09/509872 PCT/CA98/00937

19 OCTOBER 1998 (19.10.98)

OPIC
OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA

CIPO
CANADIAN INTELLECTUAL PROPERTY OFFICE

REC'D 27 OCT 1998
WIPO PCT

Bureau canadien des brevets

Certification

Canadian Patent Office

5

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,217,275**, on October 3, 1997, by **NEWBRIDGE NETWORKS CORPORATION**, assignee of Ian H. Duncan, Ken Young, Grant Hall, James Watt, Jean-Marc Ernault and Dave Watkinson, for "Multiple Internetworking Realms within an Internetworking Device".

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

S. Grégoire
Agent certificateur/Certifying Officer

October 19, 1998
Date

Abstract

An internetworking system operating over an ATM backbone. The physical internetworking devices within the system are shared to provide the internetworking functions while servicing two or more distinct and isolated user networks. This is accomplished by logically partitioning the devices into distinct sub-elements which provide all or part of the internetworking functions. These sub-elements are uniquely allocated to independent realms which are then assigned to specific user networks.

Multiple Internetworking Realms Within an Internetworking Device

Field of the Invention

This invention relates to the provision of internetworking service functions utilizing multi-protocol over ATM (MPOA) and more particularly to a system and method wherein a common backbone infrastructure is shared by several distinct user networks.

Background

Multi-protocol over ATM (MPOA) represents an important development in the communications industry in that it permits the internetworking of local area networks (LANs) over an ATM backplane. This internetworking leads to the delivery of multimedia services such as video, voice, image and data.

Currently, MPOA internetworking architectures are not capable of servicing more than one user network. Internetworking devices within the network architecture provide one or more functions related to forwarding data packets through a network. The primary keys used to control internetworking forwarding functions are network addresses. Within a particular network these network address keys must be unique for the correct operation of the forwarding functions. In many internetworking systems, in particular those based on the internet protocol, the correct operation of the forwarding functions requires the additional constraint that these network address keys are organized in an ordered hierarchy of partial address prefixes where the unique set of keys used to control the internetworking

forwarding function at different points within the network are different. In current systems, a router and bridge combination sometimes known as a ridge provides the address keys in order to forward the data packets to the proper destination.

Summary of the Invention

The purpose of the present invention is to permit the sharing of physical devices which provide the internetworking functions while servicing two or more distinct and isolated user networks. This is accomplished by logically partitioning the devices into distinct sub-elements which provide all or part of a specific internetworking function including: physical interfaces; connectivity contexts; dynamic storage and context for routing calculations; storage and context for forwarding information; storage for queuing of packets being forwarded; and the necessary storage and context of secondary elements of the internetworking forwarding functions. The sub-elements of the device are then uniquely allocated to independent realms. These independent realms are assigned to specific user networks preserving the necessary uniqueness and any local differences in the primary address keys and all other secondary information used in the correct operation of the internetworking forwarding function.

The present invention provides a distributed system built from collaborating internetworking devices and provides for large-scale internetworking services for carriers and service providers. This is known as carrier scale internetworking or SCI.

Therefore, in accordance with a basic aspect of the present invention there is provided in a system for delivering internetworking service functions utilizing internetworking devices to provide the services to two or more specific network users, the method comprising: logically partitioning the devices into sub-elements;

allocating the sub-elements to independent realms; and assigning the independent realms to the specific network users.

Brief description of the Drawings

The invention will now be described in greater detail with reference to the attached drawings wherein:

Figure 1 is a service view of a CSI system;

Figure 2 is an architectural view of a CSI system;

Figure 3 illustrates control and data traffic for internet service;

Figure 4 illustrates control and data traffic for route VPN;

Figure 5 shows one PIPI implementation;

Figure 6 is a Realm level Service Differential example

Figure 7 shows intra-realm Vnet level service differential

Figure 8 illustrates a CSI management model;

Figure 9 is a diagram of traffic and control flow to and from a PIPE

Figure 10 illustrates a simplified SCI system;

Figure 11 shows a network layer forwarding mechanism; and

Figure 12 is a PIPE instance screen

Detail Description of the Invention

CSI has a number of new terms which are described here in the hope that it will help the reader better understand the balance of this document. Refer to Figures 1 to 4 for further information on how these functions are related and interconnected in a CSI system.

1. Internetworking Services: Internet connectivity, routed VPNs, and bridged VPNs are three examples of internetworking services that a carrier may provide to customers through the CSI system.

9. Route Server: In the CSI architecture, the Route Server's main task is to generate and download the forwarding tables to the Edge and Core Forwarders. The RSes run all the required internal and external routing protocols in the CSI system to provide both default connectivity and shortcuts. The RSes are not part of the user data path.

10. Config. Or Configuration Server: In the CSI architecture, the Config Server's main tasks are: 1) to reply to requests from Edge Forwarders as to the where-abouts of their Route Servers, 2) download load configuration information to Route Servers regarding VPNs, routing protocols, and other configuration information required by the Route Server to run and 3) track the Route Servers status and activity (i.e. which RSes should be active and which ones should be on standby).

11. Shortcut VCs: These are direct SVC connections between two Edge Forwarders established for forwarding. Shortcuts are established by the EFs as a result of flow detection policies or administrative control.

12. Customer: In the CSI System, a customer is the owner of a Realm. A customer can have one or more realms.

13. Realm: The CSI System allows 3 types of realms, Routed VPN, Bridged VPN or Public Internet realms.

14. Bridged VLAN: Bridged VLAN is a way of providing Bridged VPN service. A Bridged VLAN belongs to Bridged Realm and supports multiple protocols. A Bridged VLAN operates over a set of Service Interface Groups.

15. Virtual Subnet: A Virtual Subnet is a way of providing Routed VPN service. A Virtual Subnet belongs to a Routed

Protocol Realm and supports one protocol (IP) in this description. A Virtual Subnet can be configured to operate on one or more Service Interface Groups: a Virtual subnet corresponds to one IP subnet.

16. Subnet Group: A collection of Subnets. A Subnet Group is part of the CSI Management model.

The purpose of Carrier Scale Integration (CSI) is to meet the future needs of large providers of internetworking (frame- and packet-based) services. To do so, CSI strives to meet ambitious goals in:

number of customer connection points;
number of simultaneous connected individual users;
number of simultaneous flows;
support for public and multiple private Internet packet services;
support for multiple private bridged services;
access resale with distinction of customers to a fine degree of granularity, e.g. to different end stations within a customer site;
differentiated service for both configured and dynamically detected flows;
reduction of relative management complexity;
modularity of functions, such that the CSI system works together as a whole, but functions can be replaced individually with constrained impact;
high availability;
high stability, including routing.

CSI is a distributed system built from collaborating ATM switches, route servers, access terminations, edge forwarders, default forwarders, core forwarders, a management system, and auxiliary servers. As a whole, the CSI system provides internetworking services at both the packet and frame levels. The CSI architecture defines the external interfaces between the CSI system and the outside

world and the internal interfaces between CSI components. A CSI system is expected to be managed as a whole, by or on behalf of a single service provider.

External interfaces are classified as either access interfaces or service interfaces.

Access interfaces are the interfaces over which one or more service interfaces are provided between the customer and the CSI system (e.g. STM1 UNI or 10BaseT). Access interfaces interconnect the CSI system and customer access networks, which can be any of various technologies, from a PSTN modem to a campus LAN. The concept of the access interface includes all aspects of the interface which are specific to the particular physical type of the interface as well as any interface-specific transmission protocol issues.

Access interfaces are provided by CSI components known as Access Terminations. Packets transmitted towards (and received from) the access network are encapsulated (and de-capsulated) by the access termination components. The access termination device provides all the control and auxiliary functions required by the access interfaces and transmission across them, e.g. switched-access signaling and Frame Relay LMI. Access interface does not refer to a physical interface of the access termination, but rather to a set of functions performed by the access termination. Conceptually the access interface is internal to the access termination.

Service interfaces are logical interfaces through which services are provided to the customers. A service interface is expected to carry traffic for one customer, although a customer may encompass many end systems. The control and user data flows for each service are those appropriate to the service.

Service interfaces are provided by Edge Forwarders. Edge forwarders exchange encapsulated, interface-independent PDUs (Protocol Data Unit) with the access terminations, and provide all control and auxiliary functions required by

higher layer encapsulations and control protocols such as PPP.

A service is coordinated communication between an access termination and a specific customer across a service interface, using sets of supported protocols and the management of control and user information according to those protocols. Three services are available in CSI:

1) Public Internet access service, which is managed connectivity to the public Internet.

2) Virtual private network (VPN) service, which is managed connectivity to a virtual private network. A virtual private network may include both virtual LANs (bridged connectivity) and virtual subnets (network layer connectivity).

A service enables connectivity to a Realm. A realm is a specific instance of an internet or VPN service. Within a VPN realm, there may be multiple virtual LANs for different protocol families, but only one of each. A single service interface may support multiple virtual subnet services (within a VPN realm), but only if their internet address spaces are distinct. Different PDUs from a single end station may be injected into different virtual LANs or virtual subnets.

An access interface may support more than one service interface simultaneously, but a service interface may support only one service at a time, and a service may be provided for only one realm at a time. The particular service and realm available on a particular service interface shall be controlled by configured policy, authentication and authorization.

Mechanisms for providing services and distinguishing realms are discussed later.

One aspect of service is differentiated service. Depending on the capabilities of individual access and service interfaces, and customer configuration, service may be differentiated in several ways. For example, some

traffic may be given simple priority, or weighted fair queuing schemes may be enforced. The CSI architecture is intended to allow for service differentiation at the level of individual flows, but does not require it. In some cases service differentiation might be done at the level of a whole service interface.

Finally, one or more route servers may communicate with other routing entities outside of the CSI system, for the exchange of internet routing information. From the point of view of routing, the route servers represent the CSI system to the outside world. This communication takes place at the internet layer, across an access termination or an edge forwarder.

The foundation of a CSI system is an ATM network. On this ATM network, CSI coexists with other services which might be offered, such as circuit emulation. In practice, a single ATM network may serve as all of access network, distribution fabric and transport fabric. The role of the ATM network is to provide high-speed, complete connectivity between components of a CSI system. The purpose of the nomenclature of the three fabrics is to aid in discussion.

All interfaces between the fabric and the components of a CSI system are ATM UNI (User Network Interface) interfaces.

In the CSI system, all packets within a flow of either control or user data are encapsulated using LLC (Logical Link Control) encapsulation. This permits, but does not require, multiple flows to be carried over a single VCC. Control and user data flows cannot be carried in the same VCC.

The management system provides all other CSI components with the basic configuration information they need to communicate and to establish bindings between interfaces, services and realms. Configuration information is given to each component when it becomes operational, and may also be updated at any time.

The management system itself may be made up of one or more components.

Access Terminations provide access interfaces. On the access network side they terminate data and control planes. On the CSI side of the network they provide a uniform connection mechanism and traffic stream to edge forwarders. Access terminations act as aggregation and distribution points, collecting traffic from access networks to distribute to one or more edge forwarders, and distributing traffic from one or more edge forwarders to one or more access networks. The distribution of traffic is controlled by configuration information.

The primary motivation for separating the access termination functions from the edge forwarding functions is to enable the access resale capability.

Access terminations provide limited service differentiation through traffic prioritization between interfaces. This is done under the control of the management system. Access terminations do not do any filtering or traffic shaping for incoming (i.e. from the access network) traffic. Outbound queues are FIFO queues with Random Early Drop (RED).

Edge forwarders terminate service interfaces and provide all functions related to forwarding in the CSI system, for both packets and frames. Edge forwarders are potentially the most sophisticated components in a CSI system.

While access terminations may distinguish between traffic destined to different edge forwarders, edge forwarders are responsible for more sophisticated service differentiation.

Edge forwarders receive encapsulated PDUs from access terminations and other forwarders, examine them according to rules given by the management system, categorize them, manipulate them as necessary, and forward them using rules appropriate for the realm in which the PDUs are placed. The

processing rules may lead to forwarding of either bridged frames or routed packets, in private or public nets, on a per-PDU basis.

Where the control plane of a service interface includes authentication, for example with PPP, the edge forwarder will perform preliminary authentication of users, since this may affect the distribution of traffic. Edge forwarders also provide all other functions ancillary to higher layer protocols, such as support for proxy ARP (Address Resolution Protocol) and inverse ARP, and may act as a proxy for some services such as DHCP (Dynamic Host Configuration Protocol). They may make use of other resources, such as route servers, to perform these functions. Edge forwarders represent the CSI system at the internet level, for example by responding to IP-based echo requests.

Edge forwarders inform route servers of all changes in topology concerning links to access terminations and configured links to other forwarders. Edge forwarders differentiate between flows and provide differential queuing services for flows where configured. Edge forwarders may also detect flows and create "shortcut" VCCs to other forwarders where appropriate, when allowed by configuration.

While not a basic architectural component, the concept of an Access Forwarder is used in practice. "Access Forwarder" is shorthand for a close association of the functions of Access Termination and Edge Forwarder. Architecturally the functions remain separate. In reality an access forwarder need not use the standard interface between access termination and edge forwarder.

A core forwarder is a low overhead, low functionality, possibly high speed internet-level forwarding device in the core of the CSI network, for use only by public internet services. Core forwarders are not necessary to the functioning of a CSI system, and are provided to support scalability (by making it possible to reduce the number of VCCs between edge forwarders and by offering a default

forwarding path for forwarders which cannot hold full forwarding databases). A core forwarder has no direct service interfaces and runs no routing protocols. Special features, where necessary, should be implemented in the edge forwarders and access terminations, thus allowing the core forwarder to support high speed and high capacity without high overhead. Although some end-to-end features (e.g. in Resource Reservation Protocol (RSVP) and Integrated Services) require support in all forwarders, in the core forwarder speed and capacity are far more important than feature richness.

For scaling of VPN realms, it is anticipated that it will be possible to support core forwarders which are dedicated to particular VPN realms. At this time core forwarders are intended particularly for public internet realms.

A default forwarder is essentially a more intelligent core forwarder, used in support of private realms. In private realms, edge forwarders may not have complete forwarding information. Rather than drop packets/frames while they are retrieving this information (from route servers) they forward them to the default forwarder. The default forwarder is more sophisticated than a core forwarder, in that it must take VPN policy information into account when deciding how to forward.

In the cases of both packets and frames, route servers are responsible for routing, while forwarders are responsible for forwarding. The functions of routing are explicitly separated from the functions of forwarding, in order to make it possible for individual components to do each more efficiently. Route servers are not in any user data path, and are not responsible for forwarding any user data.

Route servers are responsible for:

providing forwarders with service-related configuration information and interface bindings, and updating this information as necessary;
exchanging routing information with internal and external routing agents;
gathering information internally to keep track of internal topology;
computing forwarding databases as needed from the above information and from configured policy;
disseminating these databases to the edge and core forwarders (full tables in the public internet case; partial, full, or on-demand for private services); and
answering queries in support of other functions the forwarders may perform such as ARP.

Auxiliary servers provide support for such services as DHCP, DNS, and NTP, which run at a higher layer but are considered fundamental to normal network use. Such services are beyond the scope of the CSI architecture, but support for their functioning across the CSI system is not.

In some cases, the auxiliary server may not be directly associated with the CSI system, e.g. an exogenous RADIUS server may be used to provide AAA services, or even if it is part of the system, e.g. an internal RADIUS server, it may not be user-visible.

This category does not include "content" servers such as NetNews, web servers, electronic mail, or user directory Services.

Interfaces between CSI components support both control and user information. Interfaces occur over either "persistent" or "non-persistent" ATM SVCs. Persistent SVCs (SVC-Switched Virtual Circuit) are established per configuration, are maintained regardless of inactivity, and are re-established in the case of failure. Non-persistent SVCs are established only as needed and are released on inactivity. The particular definition of "inactivity" is a

matter for local policy, and may be part of the information obtained from the management system.

A flow of either control or user information is carried in a single VCC. Multiple flows may be carried in a single VCC, but control flows are separate from user information flows.

All configured control flows within the CSI system take place over persistent SVCs. User data flows used to provide default connectivity--that is, flows established based on configuration information and not on observed behavior of traffic or other criteria--are also carried over persistent SVCs. All other flows are carried over non-persistent SVCs.

In all cases, when a VCC is set up, ATM signaling is used to indicate the particular realm the VCC is being set up for. ATM signaling may also be used to indicate that a VCC is to be used for multiple realms, using B-LLI, B-HLI, and/or L2TP.

Each component has, as part of its basic configuration, one or more anycast ATM addresses for contacting the management system. The first connection a component establishes is with the management system over a persistent SVC. In the usual case, the management system then gives the component the information it needs to establish other default connections, and to know how to use them. These "default forwarding" connections are then established and maintained.

Specifics of internal interfaces follow.

The first connection established by any component except the management system is with the management system. This is a control interface, with no user data flow. Every component must maintain a persistent connection to the management system. In the usual case, the management system then passes configuration information to the component which the component needs in its specific situation. This policy information may include:

Access interfaces and service interfaces to be enabled.

ATM addresses and other necessary information for establishing connections with other components. Other components may include: edge forwarders, core forwarders (for all but access terminations), access terminations (for edge forwarders), and default forwarders and route servers (for all but access terminations).

Access terminations are given rules to use in determining how incoming traffic should be processed and forwarded. However, such information is not given to forwarders for their service interfaces--they obtain that information from their route servers.

What to accept connections from.

Information for route servers regarding realms, routing peers and protocols, and components for which they are responsible.

Bindings of route servers to realms and services

The management system may update a component's configuration information at any time using the interface provided by the persistent VCC.

Components may have information configured statically. Although they must connect to the management system, there is no requirement that they receive their policy information from the management system. CSI system managers may trade off the ease of central configuration management for the sake of simplicity and robustness. Hybrid schemes are possible where management information is statically configured into a component, but can be overridden by dynamically downloaded information. Protocols used for carrying information between the management system and other CSI components must be reliable.

An access termination examines incoming traffic and redistributes it to one or more edge forwarders in one or more VCCs, according to configured policy. An access termination interacts only with the management system and with one or more edge forwarders.

An access termination may bypass nearby edge forwarders and use VCCs to remote edge forwarders. This practice is known as access resale, and allows the CSI system operator to deliver traffic transparently from an access termination in one location to an edge forwarder in another location, for example to an interface to an Internet service provider.

In large-scale environments, in order to reduce the number of VCs from access terminations to edge forwarders, access terminations should support L2TP directly over AAL5 or some other scaling mechanism. Flows with different service requirements shall be carried in different L2TP tunnels.

There is no direct communication between Access Terminations. All traffic from an access termination which flows into the CSI system must flow to an edge forwarder.

A particular implementation of an access termination may allow traffic to make "hairpin turns," entering on one service interface and exiting immediately on another. Such implementations must take policy configuration into consideration. Configured policy may affect such traffic in two ways: first, with regard to the legality of the traffic flow, and second, differentiation of service.

Edge and core forwarders are responsible for establishing persistent connections to those route servers dictated by their configuration.

Route servers provide forwarders with configuration information related to service interfaces, including bindings between service interfaces and particular realms.

Route servers obtain reachability information from two sources: external routing entities (in peer networks and customer networks) and from edge and core forwarders.

The route servers obtain external reachability information through use of standard routing protocols (BGP-4 for external providers; RIPv2, OSPFv2 or BGP-4 for customer networks).

Edge forwarders send internal connectivity information (including information they obtain from access terminations) to the route servers using OSPFv2. Only topological connectivity information is sent, not information about reachable destinations. Also, ad hoc shortcut VCCs are not advertised. Finally access terminations do not appear in this topological information.

The route servers use the routing information from external sources, topology information from the forwarders, and policy information from the management system, to compute forwarding rules for each forwarder in the CSI system for which they are responsible.

They then download this forwarding information to the forwarders. As a given forwarder may participate in multiple realms, forwarding information includes at least incoming service interface, PDU characteristics such as source and destination addresses, output service interface and output queuing regime.

Route servers are also responsible for computing multicast forwarding rules for the forwarders, for use within and between realms. Multicast within bridged realms is managed following the usual mechanisms for VLANs. Since unicast forwarding rules may already include information such as incoming interface and source address, no new protocol features are required to support distribution of multicast forwarding information to the forwarders. Multicast join and leave requests are sent from the forwarders to the route servers, which then distribute the appropriate forwarding rules in response.

Finally, edge forwarders may query route servers to resolve from MAC or internetworking addresses to ATM addresses in the case of VPN traffic (both bridged and routed).

Route servers establish connections to other route servers according to configuration.

Route servers use iBGP4 to communicate external reachability information to each other. The BGP Next-Hop attribute is used to distribute the ATM address of the appropriate Edge Forwarder for external routes. This is required because the route servers may be physically separate from the forwarders.

Route servers use OSPFv2 to communicate internal topology information among themselves. Only information about configured connections is distributed between route servers. Information about dynamic, "shortcut" connections is never propagated.

Route Servers may propagate NHRP and MAC-layer address resolution queries to the next Route Server along the "default" path to the destination within that particular realm.

Given the forwarding tables delivered from the route servers, the edge and core forwarders forward IP packets as required by "Router Requirements"; this includes generating ICMP messages as required. The Forwarders also respond to ICMP Echo Messages. Further, for packets received from a customer network, the Edge Forwarders may verify that the source address is valid for the network from which the packet was received.

Edge forwarders establish connections with each other for two reasons. First, if configured to do so for a particular realm, and second, if a flow is detected and the edge forwarder considers a direct "shortcut" connection to be appropriate. In the case of a configured connection, either edge forwarder may attempt to open the connection.

Core forwarders only support the public Internet realm. Private realms (bridged or routed) always use direct connections between edge forwarders.

Edge forwarders communicate with each other using protocols appropriate to the type of realm being supported. All packets or frames are encapsulated as required by the Fabric. Data transferred as part of a routed realm are

transferred as encapsulated internetworking level packets while data transferred as part of a bridged service are transferred as MAC frames.

Shortcut connections are direct SVC connections between two Edge Forwarders, for flows which are high-volume or require specified Quality of Service (QoS) or other segregated handling. Shortcuts are established by the edge forwarders as a result of flow detection policies or administrative control. The decision of when a flow has been detected for which a shortcut connection is useful is an implementation issue.

Within a single Realm and a single QoS, multipoint-to-point VCCs can be used to reduce the number of VCCs a forwarder must support. VCCs between two forwarders may carry traffic from multiple realms. With appropriate signaling and encapsulation a single VCC may carry traffic for multiple realms as described previously.

Core forwarders forward between each other as dictated by configuration and by downloaded forwarding databases. Core forwarders do not exchange routing information, do not detect flows, and do not create dynamic "shortcut" SVCs.

With CSI, WAN internetworking service providers (e.g. ISP, telcos, IXCs (Inter Exchange), large private enterprises, etc.) can:

1. Support as a service, multiple instances of the routed Virtual Private Network service over a variety of service interfaces.
2. Support as a service, multiple instances of the bridged Virtual Private Network service over a variety of service interfaces.
3. Support as a service, multiple instances of the public Internet connectivity services over a variety of service interfaces. Note that in release Amethyst, only one instance of the Public Internet connectivity will be supported.

4. Capability to provide, support and manage all services above (routed VPNs, bridged VPNs, and Public Internet) over a single ATM network infrastructure.

5. Provide differentiated classes of service to customers for all service types.

6. Build scaleable and high bandwidth internetworks.

7. Coexist with other services offered by an ATM switch such as Newbridge Network Corporation's 36170 (Frame Relay, Cell Relay, Circuit Emulation, etc.), as well as other ATM services.

Figure 1 shows a service view of a CSI system.

The CSI network consists of the following four entities (see Figure 2):

1. A connection oriented transport fabric infrastructure provided by ATM switches.

2. Access terminations with separate or integrated edge forwarding are provided by access forwarding devices.

3. Internetworking functions (layer 3) are provided by the Route Server/s, Edge Forwarders, and Core Forwarders. Core forwarders will be optional. In Figure 2, all internetworking layer devices are shaded. All data forwarding devices are lightly shaded except the RSCP (Routing Service Control Point) is shaded differently (darker) to indicate that although the RSCP in involved in the internetworking layer is has a different function. The RSCP does not participate in the forwarding of user data but instead is responsible for running the system's routing protocols and generating forwarding tables.

4. The NMS consist of element, network, and service management systems and is responsible for managing all components of the CSI system as listed above.

.The RSCP supports routing protocols, generates forwarding tables for the edge and core forwarders, and provides address resolution as required. For scaling and availability reasons, multiple RSCPs can be deployed in a single network.

At Layer 3, the second most intelligent components in the CSI architecture are the Edge Forwarders (EFs). EFs forward IP traffic over the ATM fabric via ATM SVCs, either short or long hold SVCs depending on the type of service.

There are three types of traffic in a CSI Network:
Routing traffic -this is routing information exchanged between various routers in the network.
Control traffic - the RSCP stores control information (e.g. forwarding tables) for each of the Efs. Efs obtain this information using ATM SVCs.
Data traffic - bridged or routed PDUs being exchanged between Efs.

For routed and bridged VPN traffic, the Edge Forwarders will forward traffic to the Default Forwarder prior to a short-cut SVC being setup. Once the Edge Forwarder has set up shortcut connections across the ATM transport fabric, it will forward the traffic across the SVC and not the Default Forwarder. Access resolution is provided on demand by the RS.

For Internet traffic in CSI, the Edge Forwarders always forward the Internet traffic to an assigned CF or directly to an egress EF. The CF will then relay user data traffic based on its forwarding tables to either another CF, EF, or to an external interface (i.e. other ISP). Differentiated service for Internet traffic is possible and handled by the EFs. Edge Forwarders, with support from RSes via NHRP, will set up short-cut connections with appropriate QoS across the ATM transport fabric.

The ATM fabric provides complete data path interconnection of the CSI components. The SVC and connection oriented nature of ATM allows for cut through connections to be made on demand as required by the internetworking layer and the sophisticated QoS/TM features of ATM are ideal for mapping prioritized customer traffic to different classes of service.

The services supported include public Internet connectivity, Routed VPNs and Bridged VPNs. Each service interface must be configured for one service only although a single access interface may support multiple service interfaces.

The following sections provide a brief summary of the general functionality of each of the services.

The Routed VPN service provides Ipv4 unicast and multicast forwarding of packets received on service interfaces. Each service interface supports one or more Ipv4 subnets; the subnet prefixes need only be unique within the VPN. Routing information is exchanged between the VPN and external equipment using standard routing protocols.

The Routed VPN service will not forward traffic outside of the VPN; however nothing precludes external gateways (e.g. routers, firewalls) from providing connectivity between VPNs or between a VPN and a Public Internet service.

The Bridged VPN service provides IEEE 802.1(d) transparent bridging across a set of service interfaces, including an instance of the Spanning Tree Protocol. Each Bridged VPN can support a configurable set of protocols. Frames from a single service interface may be delivered to multiple Bridged VPNs, however the set of protocols supported by each VPN must be distinct.

The Public Internet service provides Ipv4 unicast and multicast forwarding of packets received of service interfaces. Each service interface supports one or more Ipv4 subnets; the subnet prefixes must be globally unique.

Routing information is exchanged between the Public Internet and external equipment using standard routing protocols. Subnets within the Public Internet service can be partitioned into multiple Autonomous Systems to allow multiple (routing) policy domains within a single service.

In the Example shown in Figure 3, the following interfaces and protocols are required to support public Internet services:

Both RSCP_1 and RSCP_2 support Internet routing (eBGP; iBGP and OSPF). NHRP is run on both RSCP_1 and RSCP_2 (server-server) to support EF-to-EF shortcuts as described below. Both EF_1 and EF_2 support service interfaces to Internet customers. Full forwarding tables are downloaded from RSCP_1 to EF_1 and RSCP_2 to EF_2 via the Table Download protocol .

Shortcut data paths for higher CoS may be established for Internet services between EF_1 and EF_2 based on administration control or configured policies in the EFs. A client is run in the EFs to perform address resolutions.

In the example of Figure 4, the following interfaces and protocols are required to support Virtual Subnet services:

EF_1 supports R-VPN_A Service Interfaces using RIP as the routing protocol and VPN-B Service Interfaces with OSPF as the routing protocol. EF_2 supports R-VPN_A and R-VPN C running RIP and R-VPN B running OSPF.

For VPN_A, an instance of RIP will run between RSCP_1 and EF_1 VPN_A attached devices and similarly between RSCP_2 and EF_2 VPN_A attached devices. For full reachability, an instant of RIP associated with VPN_A operates between RSCP_1 and RSCP_2.

For VPN_B, an instance of OSPF will run between RSCP_1 and EF_1 VPN_B attached devices and an instant of OSPF between RSCP_2 and EF_2 VPN_B attached devices. To fully manage VPN_B across the two RSCPs, an instant of OSPF associated with VPN_B is run between RSCP_1 and RSCP_2.

For VPN_C, an instance of RIP will run between RSCP_2 and EF_2 VPN_C attached devices.

Shortcut data paths are established between EF_1 and EF_2 for all Unicast data traffic. A client is run in the EFs to perform address resolutions for shortcuts via the RSCPs. NHRP is run on both RSCP_ 1 and RSCP_2 to support EF-to-EF shortcuts. EFs maintain a cache of most frequent connections (to minimize EF-RSCP activity) and connections

are based on resilient SVCs (to minimize SVC set-up/tear-down).

Directed broadcast and multicast traffic is forwarded to the RSCP's internal DF as shown in Figure 4. Using direct p-to-mp connections the DF is responsible for forwarding the traffic to the egress EFs. The internal DF is also used for providing unicast forwarding for VPNs during the detection and set-up time of short-cut connections (SVC)..

Table 1 summarizes the performance of a CSI System. Unless otherwise noted, the numbers shown are the minimum supported performance level under any condition.

The latency numbers quoted for the PIPE should be valid for situations where the total traffic being forwarded by the PIPE is less than the backplane bandwidth available to the PIPE, and when all traffic is treated at the same priority level. Once the backplane bandwidth is exceeded, latency becomes a function of PIPE output queue depth at any given instant. In the case of multiple COS, the latency numbers quoted should be valid for the highest-priority output queue. Note 1: Latency targets assume no congestion in the network. To calculate the typical latency of a packet traversing a CSI network, simply sum the individual latencies. For example, if a packet goes from a PIPE to a RIDGE, and traversing 5 switches in the process, the typical latency will be P2 + P3 * (5 * P5

	Criteria	Phase 1 Target	Phase 2 Target
P1	System Restart time (cold start)	15 minutes	15 minutes
P2	Packet latency, PIPE (128 byte packet, typical)	40 µs	40 µs
P3	Packet latency, Ridge (128 byte packet, typical)	100 µs	100 µs
P4	Packet latency, 36170 (typical) (note 1)	35 ms	35 ms
P5	Shortcut path setup time, once a flow has been detected (typical)	20 ms	20 ms
P6	RSCP Integral Default Forwarder unicast forwarding	10,000 pps	50,000 pps
Pn	RSCP Integral Default Forwarder multicast forwarding	1,000 pps	50,000 pps
P7	Yellow Ridge IP Unicast Packet Forwarding Rate (packet size = 128 bytes)	84,400 pps	84,400 pps
P8	Orange Ridge IP Unicast Packet Forwarding Rate (packet size = 128 bytes)	84,400 pps	84,400 pps
P9	Red Ridge IP Unicast Packet Forwarding Rate (packet size = 128 bytes)	TBD (has not been characterized yet)	TBD (has not been characterized yet)
P10	PIPE IP Unicast Forwarding Rate (packet size = 128 bytes)	118,000	118,000
P11	PIPE IP Multicast Forwarding Rate (packet size = 128 bytes)	TBD	TBD
P12	ICMP request handling on RSCP	10 per second	10 per second

P1 3	ARP requests handled per RSCP	500 per second	500 per second
P1 4	IGMP requests handled per RSCP	TBD per second	TBD per second
P1 5	OSPF updates absorbed per RSCP	TBD per second	TBD per second
P1 6	BGP-4 updates absorbed per RSCP	TBD per second	TBD per second
P1 7	Maximum Service Outage During RSCP Activity Switch	2 minutes	2 minutes
P1 8	Maximum service outage during PIPE activity switch	20 seconds	20 seconds
P1 9	RSCP Max Route Change processing rate (routes)	25k per second	25k per second
P2 0	Forwarding Table Download Rate (routes)	1000 per second	1000 per second
P2 1	Number of SVCs per second originating from PIPE	100 calls/seco nd	100 calls/seco nd
P2 2	Number of Address Resolution Requests per PIPE	800 per second	800 per second
P2 3	Number of Address Resolution Requests per Ridge	50 per second	50 per second
P2 4	Number of Address Resolution Requests per Tigris	TBD	TBD
P2 5	Multicast Forwarding Rate (Multicast Server)	50,000 pps	50,000 pps
P2 6	Service Unit Restart Time	TBD	TBD
P2 7	RSCP Restart Time	TBD	TBD

Table 1 CSI Performance Summary

The Packet Internetworking Processing Engine (PIPE) provides a high-fanout Edge Forwarder as a 36170 UCS card. This engine is used to forward IP traffic delivered to the system on FR, PPP or ATM interfaces (see Figure 5) In the case of RF or PPP traffic, the sessions must first traverse a Frame Relay card in the 36170, however this card can be in a different shelf or system from the PIPE.

The PIPE, based on the Extended Processing Engine Platform, provides the following instructions:

a) automatic download of configuration information from the Configuration Server,

b) initiation of SVCs as required to provide connectivity,

c) termination of PPP sessions and FR connections,

d) support for C10 independent forwarding contexts with a limit of C18 total forwarding entries per PIPE,

e) obtains forwarding information from a Route server,

f) packet classification and output queue selection in support of system-level traffic management policing,

g) transparent bridging in support of the Bridged VPN service,

h)IP unicast and multicast forwarding in support of the VPN and Public Internet services, and

i) N+1 redundancy

The ATM fabric provides interconnection of the CSI components for both control and user-data traffic. As shown in Figure 2, each component of the CSI System is connected to the ATM fabric; connectivity between components uses ATM Virtual Channel Connections (VCCs).

Most inter-component SVCs are "resilient, long hold time" SVCs, i.e. they are (re)established on component restart. On-demand SVCs are only used to provide shortcuts for the VPN service. The "resilient" nature of the SVCs indicates that the component that originally initiated an SVC will persistently attempt to re-establish the SVC if it is ever cleared by the network. The interval between such

re-establishment attempts is subject to an exponential backoff.

The generation of SVC setups by a component is rate-limited.

There are three primary categories of inter-component connectivity; these are described in the sections that follow.

The CSI System uses three set of VCCs for connectivity in the control plane:

a) from an Edge Forwarder to the Configuration Server for configuration information download

b) from the Edge Forwarder to the Route Server for basic control function and on-demand address resolution for VPN services and

c) from the Route Server to all of the Edge Forwarders for distribution of forwarding table information in support of the Public Internet service and basic control.

A unicast SVC is established from the Edge Forwarder to the RS/CS for registration and cache management. The RS/CS then establishes a LAN Control SVC back to the Edge Forwarder over which configuration is downloaded with guaranteed delivery. The RS/CS also adds the Edge Forwarder as a leaf of P2MP SVCs, one for each VPN.

Traffic descriptors for all types of connections, except the RS SVCs, are configurable. The non-service interface connections are only configurable on a per-category per-realm basis.

The defaults for all data connections (service interfaces, short-cuts, default forwarder connections, etc.) are UBR, PIR = line_rate, MIR = 0 bps.

The defaults for all control connections (to control server and route server from PIPE) are: nrtVBR, PIR = line_rate, SIR = TBD, MBS = 32 cells, CDVT = 250(s.

Each Edge Forwarder obtains from the Configuration Server the ATM addresses of all Edge Forwarders involved in

Public Internet traffic forwarding, or of a Core Forwarder, to which it maintains ATM connectivity.

The Edge Forwarder maintains a VCC to each Edge Forwarder and/or Core Forwarder for each class of service; this VCC is established upon restart and/or (re)configuration.

Each Edge Forwarder obtains from the Configuration Server the ATM address of at least one Default Forwarder to which it maintains ATM connectivity. The Configuration information supplied by the Configuration Server results from the configuration of the system.

The Edge Forwarder maintains a VCC to each Default Forwarder for each class of service; this VCC is established upon restart and/or (re)configuration. If the VCC is released by the network, the Edge Forwarder persistently attempts to re-establish the VCC.

In addition to the base connectivity, an Edge Forwarder will set up a new short-cut VCC or re-use an existing shortcut VCC when it detects a flow that requires a class of service for which there is no short-cut VCC. Short-cut VCCs are disestablished, using a distinct clearing cause, when the VCC has been idle for some period of time.

Traffic Management is handled independently on a per-connection basis. There are two major types of connections in CSI, Service Interfaces and the set of SVCs comprising the CSI Core. Each connection needs the standard ATM Traffic Descriptor plus additional parameters comprising the packet-level traffic information. Note that control and routing traffic gets priority over the data traffic.

Two classes of service are provided by the CSI system. These are :

Best Effort (no guarantees for either delay or packet loss)

Better Effort

Different levels of service can be offered to different Realms in a CSI system. The Realm differentiation is achieved by configuring different sets of ATM traffic

parameters to apply to the ATM fabric CVCs for each Realm (See Figure 6). This differentiation applies only to EF-EF SVCs. There is no differentiation on the EF-RS and EF-CONS SVCs that are shared between the different Realms. In fact, there are two different SVCs per EF-EF pair, in order to allow intra-Realm service differentiation.

The Vnet level service differentiation allows prioritization of the traffic inside a given Realm. Each Vnet can be configured with the standard Best Effort Class of service or with the higher Better Effort COS. The traffic received from or transmitted to a Vnet configured with Better Effort gets ths Better Effort Class of Service. (See Figure 7.)

This same principle applies in the same way in a VPN Realm, to traffic routed or bridged between virtual subnets or VLANS or in a Public Internet Realm to traffic routed between subnets.

Effective, Better Effort COS is delivered when required by the use of separate transmission queues on the Service Interfaces of the EFs or separate EF-EF SVC over ATM fabric for each COS and each Realm.

The role of the Packet Classification is to determine the COS for each packet in the CSI System. The Packet Classification is performed on each Packet Receive Interface of the Edge Forwarders. The RS does not perform any Packet Classification in this version.

There are three different COS, from the highest to the lowest priority,

a) Contol Traffic,

b) User Data Better Effort,

c) User Data Best Effort.

In general, higher priority means lower delay and lower packet loss rate.

The Control Traffic gets the highest priority in the system to provide immunity from data-plan congestion. The Control Traffic includes,

- ARM and CCP protocols
- Routing protocols: RIP, OSPF, and BGP
- Spanning Tree BPDUs.

The Best vs. Better Classification of Data Traffic requires explicit configuration at the Vnet level.

Packets received from an SI and forwarded to the RS with BME encapsulation can get 'Control Traffic' or Best Effort COS. Routing protocols and Spanning Tree Protocol gets Control Traffic COS. Every other User Data packet falling in one of the exceptions cases gets the standard Best Effort.

Packets received from an SI and forwarded directly to another EF or internally to another SI of the receving EF get Best Effort or Better Effort COS. The general principle is that the COS is configured per VNet on the NMS. Each Vnet is configured with Best Effort or Better Effort COS. Each forwarded packet gets the higher COS configured on the source and destination Vnets.

COS (packet) = MAX.(COS(source VNet), COS(destination Vnet))

The term VNet refers to,

- Virtual Subnet for routed traffic in VPN service.
- VLAN for Bridged traffic in VPN,
- Subnet for Public Internet case.

In the Public Internet case, there is a one to one mapping between Subnet and SI, so that the 'per Subnet' COS configuration is equivalent to a 'Per SI' configuration.

The following exceptions apply to IP routed traffic.

- There is no COS differentiation between different Subnets behind a router. For traffic received from (resp. transmitted to) IP Hosts behind a router, the source (resp. Destination) VNet that is taken in account is the VNet where the router is admitted.

- In case of IP multinetting on a given SI, the different Virtual Subnets appearing on a single SI must be

configured with the same COS. This is because there are some cases when the EF cannot determine the source VNet for traffic received from hosts behind a router. This restriction does not apply to different VLANs configured on a single SI.

- In multiple RS architectures. Because COS parameters cannot be exchanged between RSs, when an EF transmits a packet to an EF that belongs to another RS domain, this packet gets the COS of the Source Vnet.

Packets received from another EF on a EF-EF SVC can get Best Effort of Better Effort COS. The packet classification is similar to the Ingress EF case, but, as there is no Source Look-UP on Egress, the source VNet COS cannot be taken into account.

It is replaced by the COS of the EF-EF where the packet is received. COS is associated with each EF-EF SVC.

COS (packet)=MAX.(COS(Receiving EF-EF SVC), COS (destination VNet))

Every ARM and CCP protocol packet received on the EF-CONS SVC or ont he different EF-RS SVCs gets 'Control Traffic' COS. Notice that this is only an internal EF classification as this type of packet is not sent to any SI.

Packets received on the LAN Broadcast SVC from the RS with BME encapsulation can get 'Control Traffic' or Best Effort COS for transmission to an SI. Routing protocols and Spanning Tree Protocol gets Control Traffic COS. Every User Data packet falling in one of the exceptions cases gets the standard Best Effort.

The table below summarizes the distribution of the User Data Packet Classification on the CSI components.

NMS	-Configuration of the COS for each VNet.

RS	- Distribution of the COS configuration to the Forwarders. - Although the COS is configured at the VNet level on the NMS, it is also stored at the Cache Entry level in ARM messages and EF Forwarding table. This ensures evolutivity to better granularity. - The RS gets from the VNM the COS configured with each VNet of its own domain and uses this information to determine the COS to assign to each Cache Entry downloaded to the Efs. - A COS parameter is thus configured with each (MAC Station, Protocol Family) for Bridged traffic and to each IP Host for IP routed traffic.
EF	- Support of the COS parameter in Forwarding Table Entries. - Packet classification on Ingress and Egress Forwarding.

Table 3-1 Packet Classification on the CSI components

The role of this function is to offer to each packet the level of priority requested as the result of the Packet Classification function. This function is implemented in each Packet Output Queuing point in the Edge Forwarders. The Rs does not perform any kind of Packet Traffic Management.

The EFs implements two separate output Queues for each SI.

Packets classified in User Data Best Effort are placed in the Low Priority Queue. Packets classified in User Data Better Effort COS and Control Traffic COS are placed in the High Priority Queue.

The High priority Queue needs to be completely exhausted prior to the Low Priority Queue being processed.

Within the High Priority Queue, lower packet loss rate is ensured to the Control Traffic COS through the use of two different Discards threshold: one threshold for User Data Better Effort COS and one higher threshold for Control Traffic COS.

A single Discard threshold is used for the Low priority Queue.

In Summary, three Discard thresholds are defined for each SI, one threshold per COS. A simple tail of Queue discard is performed for each COS: the arriving packet is discarded if the threshold is reached. Three Discard statistic counters are associated with each SI, one for each COS.

In the Ridge case, each SI is a separate Ethernet port and there is no need for a Queue servicing algorithm across the SIs.

There are multiple separate transmission Queues on the Ridge ATM port, corresponding to different transmission priorities.

Figure 8 is an illustration of the CSI management model. As this figure shows customers can have one or more realms. Each realm will have a type associated with it, one of bridged VPN, routed VPN or public Internet access. A bridged realm can have one or more VLANs associated with it. A routed or public Internet access realm can have one or more subnets or subnet groups associated with it. With each subnet group there is a set of subnets.

In addition to the common features listed above, the following features are provided for the Public Internet service:

i) The maximum number of prefixes (routes) per Public Internet Services is C4.

ii) The CSI system uses External BGP (eBGP) to exchange routing information with peers.

iii) The CSI system can use iBGP, eBGP, OSPF or RIPv2 to exchange routing information with customers; alternatively

it can use static information about what is reachable on the customer end of a service interface.

iv) The CSI system uses Internal BGP (iBGP) to synchronize the externally-obtained reachability across the Route Servers.

v) The CSI system uses OSPF and/or static routes to manage the internal topology, i.e. the pre-defined reachability between Edge Forwarders, of the components that support the Public Internet Service.

vi) The CSI system combines both the internal and external topology information while building the forwarding table.

vii) Support for multiple autonomous systems within a single Public Internet service.

viii) Unnumbered interfaces are supported.

The 36170 Access Forwarder—the Packet Internetworking Processing Engine (or PIPE: is an element developed for the Carrier Scale Internetworking System (CSI).

The following covers all functionality relating to: the termination of PPP and FR connections for carrying network traffic between the PIPE and the access interface cards; and the internetworking forwarding services necessary to process the network traffic to and from the peers on the PPP and FR service interfaces. The hardware used to support the PIPE is a 36170 control card

The PIPE is used within 36170 networks as an element of the Carrier Scale Internetworking System. The primary function of the PIPE is to provide packet internetworking (layer 3+) service boundary for a wide range of low to medium speed 36170 access interfaces

The Packet Internetworking Processing Engine provides the following primary functions:

Fn1: UCS behavior

Fn2: Virtual Connection support

Fn3: Packet forwarding

Fn4: PPP/ATM link termination

Fn5: 802.1(d) Spanning Tree Protocol (STP)

Fn6: Realm identity & network address assignment

Fn7: "MPOA" client

Fn8: Redundancy

Within the CSI system the PIPE provides the routed (layer 3) and bridged (layer 2) forwarding services for various physical Access Interfaces across a range of 36170 packet and cell interface cards. Together the PIPE and its associated Access Interfaces create a high fan-out Edge Forwarder. The two network elements described in detail herein are the PIPE card and the Access Termination/Access Interfaces as provided by the various packet and cell cards.

The CSI system is designed to give a network operator facilities to provide a range of internetworking services to customers. Figure 9 provides a simplified schematic diagram of the flows of traffic and control data to and from the PIPE. The two boxes at the left and right represent Customer Equipment (CE) that require internetworking connectivity. Typically these boxes are routers and/or bridges with some form of WAN interface which would be connected into the CSI system.

In a simple application CE, might be a router with: an Ethernet interface servicing a customer LAN; and a T1 interface providing the connection into the CSI system. The Access Termination (AT) on the 36170 would be a T1 port on a UFR card. There are two internetworking packet encapsulations which can be supported in this case. The first is Frame Relay and the second is PPP. In both cases the UFR card provides an Access Interface onto an ATM VC which connects to the PIPE across the 36170 ATM fabric. And again in both cases the PIPE provides all the necessary

functions to process the encapsulations and forward the internetworking packets flowing to and from CE,.

The Route Server (RS) provides the control information about forwarding so the PIPE can select the correct paths for delivering packets. The Default Forwarder (DF) and Edge Forwarder (EF) elements together provide the internetworking path between the PIPE and CE2. The EF element could be either another PIPE/AT pair a VIVID Ridge or a Tigris In the simple case packets will flow to and from CE, though a path that goes from the PIPE up to the DF and on through the EF to CE2. When it has been determined either automatically or through configuration that traffic between CE, and CE2 (or more correctly traffic between the PIPE and the EF) is significant enough to require a more direct path a "short-cut" connection is established directly between the PIPE and EF. Once the "short-cut" is set up traffic between CE, and CE2 will flow over the "short-cut" bypassing the DF.

In the "Public Internet" service case the connection providing the direct path between the PIPE and EF is configured administratively as a fixed link. This connection is established within the system at initialization when the components elements involved reach the full operational state and is maintained continuously.

Figure 10 provides a more complete picture of a small but typical system, showing the relationships between various the elements of the CSI application. There are a few elements, the Configuration Server (CS) and the Core Forwarder (CF), added that complete the system along with a few PIPEs, ATs and RSs illustrating the modular nature of the CSI system. The CS provides the PIPEs and other elements in the system with the details about connections and other parameters necessary to bring the system to an operational state. The CF provides a function similar to the default

forwarder in networks where the traffic characteristics requires very high capacity default forwarding paths, e.g. services providing access to the Public Internet

Figure 10 also illustrates how a small but typical CSI system could be used by a network operator to provide a mix of services to various customers while maintaining necessary partitioning of control information and traffic load.

The PIPE does not provide any external physical ports, consequently ports are not physical but are simply implementation abstractions.

The EPEC card hosting the PIPE card can be reset through system software as a maintenance function or mode reconfiguration from NMTI. Software resets will tear down all active circuits and PPP connections immediately.

The PIPE has its primary physical attachment to the network fabric via the Newbridge ATM interface to the 36170 backplane. Connections into the PIPE for various the functions detailed below are provided via PVCs, SVCs and SPVCs.

Aggregates to the CSI core are supported on conventional multiprotocol VC terminations and are either statically assigned or dynamically bound SVCs using the "MPOA" client function (Fn7). Frame Relay, PPP or ATM circuits providing network layer encapsulation services are terminated on the PIPE as PVCs or SPVCs, using this same termination function, via the FRF.8 Inter-Working Unit on the various supported 36170 frame relay interface cards. PPP packets are transferred between the PIPE and the supported 36170 interface cards using PVCs or SPVCs over a PPP/ATM transparent HDLC encapsulation.

The following table shows all of the connection types supported on the PIPE:

Connection Type	Supported Cards	PVCs	SPVCs	SVCs
Frame Relay Service Interfaces	All Frame Relay Cards	Yes	Yes *	No
ATM Service Interfaces	All Cell Relay Cards	Yes	Yes	No
PPP Service Interfaces	All Frame Relay Cards	Yes	Yes *	No
PPP over FR Service Interfaces	All Frame Relay Cards	Yes	Yes *	No
PPP over ATM Service Interfaces	All Cell Relay Cards	Yes	Yes	No
Short-cut Paths between Edge Forwarders	All Cell Relay Cards	No	No	Yes
Fixed Link Paths between Edge Forwarders	All Cell Relay Cards	No	No	Yes
Control Connections to Route and Config Servers	All Cell Relay Cards	No	No	Yes

* This is only supported if all NNI Cards are Cell Relay. Frame Relay and PPP SPVCs are only supported over the Cell Relay SVC infrastructure in this release.

Table 3-1 Connection Types supported by the PIPE

Several SVC connections must be maintained continuously to provide proper functioning of the CSI system. If one of these persistent connections is released, a call attempt is made, again to the same destination address or, if more than one destination address is available, the full set of possible destinations. The call attempts are made with an exponential backoff on failure with the initial time between attempts starting at a base interval (e.g. 1 second), after 8 attempts it does not increase further (e.g. starting at 1 second the final backoff interval will be just over a minute – 64 seconds) but the PIPE may continue to attempt the call indefinitely. The behavior if the 8th and final attempt fails is particular to the type of connection, some will persist indefinitely and others will stop at the 8th attempt and raise an alarm. The PIPE is responsible for determining if any information preserved over the reconnect has changed during the outage and reacting to these changes.

• Transport services and applications above IP (and other best-effort layer 3 protocols) are sensitive to cell loss, and the upper-layer windowing protocols will tend to drive loads to the threshold of congestion for the network, however, early packet discard schemes are available which reduce the effect of congestion in the ATM fabric and provide improved feedback to properly behaving windowing mechanisms. A simple form of ATM traffic shaping is

performed on the PIPE on a per-VC basis for traffic toward the backplane. Traffic Policing is unnecessary for the PIPE as it is a trusted UNI device. The operator can define the traffic contracts for specific categories of VCs initiated from the PIPE. These categories are:

1) Connections to the Configuration Servers;
2) Connections to the Route Servers; and
3) Short-cut connections to other Access Forwarders.

The service interface traffic parameters can be any valid selection as specified in the traffic management documents referred to above. It is intended that a network management platform support a profile mechanism for service interfaces. This reduces the amount of configuration required for each service interface. This is solely a management construct. Each service interface at the PIPE is controllable separately.

The PIPE implements services within ATM AAL5 encapsulation are compatible with the multiprotocol LLC/SNAP encapsulation. This provides IP/ATM, transparent bridging over ATM and PPP/ATM functions. This are used to provide two features within the CSI System. The first is to provide the termination for connections provided on the Access Interfaces of the CSI system including:

1. access over native ATM services;
2. internetworking with external Frame Relay attached network layer devices via the FRF.8 service IWU; and
3. PPP attached devices as provided on the various 36170 FR interface cards.

The second is to provide the connectivity over short-cuts and statically configured VC paths across the core fabric to other networking elements in the CSI System.

The basic network layer forwarding mechanism is common to both bridged and routed networks. The model for this mechanism is illustrated in Fig. 11.

The PIPE supports a fixed number of realms. The realms on the PIPE are autonomous such that each realm has its own set of FIBs and no forwarding/routing information or other state is shared between the realms. This allows the realms to have non-unique address spaces if required and, more generally, isolates the realms from one another with respect to network address assignments.

For any particular Realm, one of the aggregate interfaces will likely be configured as a connection to the default forwarder. Forwarding information about the other interfaces is either configured satirically through one of the management interfaces or via "MPOA" (Fn7). Finally, the PIB will be updated automatically with the new link-local forwarding information when PPP, Bridged or IP/ATM and Bridged or IP/FR-ATM Service Interfaces are initiated or when Service Interface is disabled (either administratively or when the underlying connection closes).

An essential element of packet forwarding on the PIPE is the process used for discarding traffic when queues reach an overflow state. The PIPE provides a two discard disciplines which are applied to the output queues. The first is a variant of Random Early Discard and the second is simple head-drop discard. The output queuing control is provided per service interface with a default setting of RED enabled.

With RED turned on, as the output queue approaches an overflow state, packets are discarded with a pseudo-random selection of the packets to discard exponentially weighted

towards the earliest packets arriving. This is a simplified description of RED.
When RED is disabled, the transmit queues operate in a simple FIFO discipline with discards performed at the tail of the queue as it reaches an overflow state.

In the extreme case where overflow occurs on input the PIPE card discards on the tail of the input queue as new packets arrive.

In addition to the packet output queuing controls and ATM level traffic descriptors applied against a connection (an access service interface, a connection to a default/core forwarder, or a short-cut path), the following additional network-level traffic management parameters.
The class of service (COS) can be one of the following values:
1. Best Effort - There are no guarantees of packet loss nor delay in the PIPE;
2. Better Effort - There is at least a 10-7 probability of packet loss within the PIPE and the packet delay is also less than for the "Best Effort" class of service; or
3. Mixed Effort - elements and attributes of each packet determine whether a best effort or a better effort class of service is to be chosen.

For packets flowing from ingress connection A to egress connection B, connection B has one best effort queue and one better effort queue. If connection A is specified to have a best effort CoS, then the best effort queue is used. If connection A has a better effort CoS, then the better effort queue is used. If connection A has a mixed effort CoS, then both queues are used. Traffic is shaped out from the aggregate of these two queues at a rate which matches the ATM traffic descriptor. The packets are emitted at a packet rate approximating the bit rates specified in the traffic

descriptors. When a packet is allowed to be transmitted according to the shaping rule table segmentation into cells occurs, the cells are sent back-to-back across the backplane of the 36170. The MBS (when shaping to the SIR) or CDVT (when shaping to the PIR) values must be chosen to ensure that the traffic contract is maintained using this form of shaping.

For each service category, the traffic is shaped according to the following traffic descriptor values:

Service Category	The rate at which packets are emitted
UBR	PIR (peak)
nrtVBR	SIR (sustained)
rtVBR	SIR (sustained)
CBR	PIR (peak)

Table 3-2 Packet Transmit Rate for different Service Categories

For routed and bridged VPNs which the "MPOA" client lookup cache management function, the packet forwarding function applies a flow detection mechanism on source-destination sets which are not currently in the cache. This mechanism monitors the traffic for the new source-destination pair and identifies the traffic as a flow when the traffic reaches a rate of at least M packets in N seconds. The default values are 4 packets in 10 seconds. Only when a flow is detected does the "MPOA" client establish a short-cut path.

Ip forwarding is the internetworking layer applied to each packet received on an IP routed service interface. This includes applying error checking rules and policy filtering, determining what to do with the packet in terms of the next-hop to its ultimate destination and finally queuing the packet for output or possible local delivery. Although Routed VPNs and Internet Access appear on surface to be significantly different features, when examining the PIPE IP

forwarding function those differences are mostly superficial. Routed VPNs tend to have a smaller set of address prefixes which change over time driven by supporting flow detection and consequently triggering "short-cuts". Internet Access typically requires a very large set of address prefixes which will change over time mostly based on updates provided by the route server via the Full Table Download function and the set of active interfaces will be relatively constant.

The IP forwarding function on the PIPE provides for support for processing IP packets which are forwarded in and out of service interfaces which are operating using the LLC/SNAP bridged encapsulation. This function provides the necessary ARP capabilities to bind and maintain MAC addresses for the IP hosts on the remote LAN segment. This function is not supported for PPP bridged interfaces.

). The IP forwarding mechanism (IFM) works by using various layer 3 information within each packet (along with information about which interface the packet arrived on) and switches packet traffic between the various PPP and IP/ATM links.

The following is a simplified description of the IFM with the terminology aligned to CSI:

1) the forwarder receives the IP packet (plus other details) from the link layer;

2) the forwarder validates the IP header;

3) the forwarder performs processing of most of any IP options;

4) the forwarder examines the destination IP address in the IP header against the FIB and assuming it satisfies basic requirements for forwarding;

5) the address of next hop for the packet (and the correct output interface) is determined;

6) the source address is tested for validity and any administrative constraints are applied;
7) the forwarder decrements TTL and then tests for expire;
8) the forwarder performs processing of any IP options which could not be completed in step 3;
9) the forwarder performs any necessary IP fragmentation;
10) the forwarder determines the link layer address of the next hop for the packet; and
11) finally the forwarder queues the packet for delivery on the interface out to the next hop.

For directed diagnostic an IP forwarding table dump is provided to verify the operational state of the FIBs The PIPE supports bridge forwarding within designated bridged VPNs. Bridging is available between service interfaces which belong to the same VLAN and protocol family(s). Bridge forwarding on the PIPE can be characterized as half bridging since it is connected to another bridge via a point-to-point link.

Diagnostics on the PIPE for Bridge Forwarding include a bridge table dump and view of the current state configuration of spanning tree. This forwarding table dump and STP view matches the elements contained in the Bridge MIB.

The bridging function on the PIPE card is determined by the configuration information sent to it by the RS. This configuration includes the definition of VPNs, VLANs and the services they offer. A service interface or set of service interfaces can only be bound to a VLAN or set of VLANs. With this information configured on the PIPE the bridge function only forwards traffic between service interfaces in the same VLAN. In this way, traffic is forwarded to only a subset of service interfaces.

The Bridging Algorithm used for the PIPE follows the standard defined in IEEE 802.1. The following functions are performed by the PIPE as part of its bridging role

1) Bridge packets from one Bridging interface to another;
2) Learning and Cache Management; and
3) Filter packets to prevent loops (informed by Fn7, the 802.1 (d) Spanning Tree Protocol).

The first function is the basic relay of packets from one end station to another on a different interface. The basic process is:
1) Bridged Packets are received by the PIPE;
2) The MAC address and service interface association of the sender are recorded in the PIPE's cache;
3) The Destination MAC contained in the packet is examined and matched to an entry in the PIPE's existing cache;
4) If an entry exists (the cache contains permanent entries for the reserved MAC broadcast and multicast addresses), the packet is passed out the associated output interface (for the broadcast/multicast entries this is the DF which then provides the correct flooding);
5) If an entry does not exist, a message is sent to the "MPOA" client function (Fn8) which will attempt to get a resolution for Destination MAC;
6) If the Destination MAC is resolved, the packet is passed out the associated service interface (in same manner as step 4); otherwise
7) The packet is discarded.
The second function is MAC address learning and cache management. When packets are received by the PIPE, a record of the source MAC address and its related service interface is kept in a cache. This cache allows the PIPE to easily look up the relationship between the source and destination identified in the packet. If the configuration for the source and destination match, the packet is forwarded to the

appropriate service interface. However, if the configuration does not match, the packet is discarded or checked for special handling, in the case of the RS, which is required to communicate with all stations. The size of the cache, however, is not infinite so an aging mechanism is required to maintain a set of recently used records for source and destination to service interface/VLAN mappings. The aging function determines whether a cache entry has been used recently. If the entry has been used it is refreshed and maintained in the cache. If has not been used, the entry is deleted to make room for new cache entries.

The PIPE card will generate billing records every fifteen minutes using the same format as using by 36170 SVC records. Information will be provided in the records for transmitted packets, received packets, transmitted bytes, received bytes. Records will also be created when the PVC is disconnected. This will provide the data for the final portion of a fifteen minute interval for which the PVC was connected.

The Point-to-Point Protocol (PPP) provides an interoperable method for communicating multi-protocol network datagrams. The PIPE provides for the PPP termination of standard bit-synchronous PPP over HDLC connections into the 36170 CSI system by internetworking with the transparent HDLC frame forwarding function on 36170 FR cards which has an optional mode for providing an internetworking service which supports conversion of PPP packets to and from the PPP over AAL5 encapsulation. This function is intended to support the "leased-line" mode of operation for permanent IP services, for example T1/E1 ISP customer "feeds
LCP options are set by the network management entities through the service configuration for a particular realm and loaded through the "MPOA" Configuration.

. The PIPE provides for static configuration of the authentication control information including the shared secrets used within the protocol. These are configurable via the network management entities and is normally loaded through the "MPOA" Configuration Server
The IP Control Protocol is used on fully established and authenticated PPP links to negotiate the IP address at each end of the PPP link and to negotiate VJ TCP/IP header compression The peer's IP address can be assigned or discovered and verified with this protocol, dependent on how the link has been configured to negotiate this option. By default, address assignment for the link peer and link local assignment from the peer are both disabled on the PIPE.

Van Jacobson TCP/IP header compression, an option that can be negotiated in IPCP can reduce a standard 40 byte TCP/IP header to variable size header between 3 and 16 bytes for most of the TCP packets transmitted over a PPP connection. VJ header compression and decompression is a function supported on the PIPE. By default, it is disabled but it can be enabled on individual PPP service interfaces through the management interfaces. The use of VJ header compression does have an impact on performance and other resources in the PIPE. In addition, depending on the nature of traffic flowing across the link and the number of "VJ slots" assigned to it may provide little or no compression.

The IETF standard PPP network control protocol (NCP) for bridging, the Bridge Control Protocol is used on fully established and authenticated PPP links terminating on the PIPE to negotiate the operation of transparent bridging of 802.3 LAN traffic. Until PPP has reached the Network Layer and BCP is fully negotiated, bridged data packets will be discarded by the PIPE.

Transparent bridging is accomplished by negotiating the following BCP options:

BCP Option	Type	Description	Length	Default Value
MAC-Support negotiation	3	MAC type traffic supported Possible values : 1=802.3Ethernet only	3	802.3
Tinygram-Compression	4	Compression of a small PDU that has padding provided the PDU is smaler than the minimum PDU size and has a LAN Frame Checksum Possible values: 1=enabled,2=disabled	3	1
Mac-Address	6	Ability to have MAC Address announced or assigned	8	
Spanning-Tree-Protocol negotiation	7	Negotiate version odf STP Possible Values: 0=NULL, 1=802.1(d)	3	802.1(d)

The CSI system provides no support for the LAN-Identification option and, because there is no requirement, there is no support for options related to source-route bridging or proprietary Spanning Tree Protocols.

The Internetworking Realms on the PIPE provide an abstraction for organizing related service interfaces; the lower layer PPP and FR access ATM VC interfaces and associated aggregate interfaces into the core networks; and the addressing information of external network services required for normal operation

The PIPE supports a fixed number of independent realms and a fixed number of service interfaces. These interfaces are distributed across realms ensuring that each realm will have a fixed number of interfaces. For example, a PIPE supporting a maximum of 500 interfaces and 5 realms might be configured to handle 3 routed IP realms, 1 with 200 interfaces and 2 with 50 interfaces, and 2 bridged realms each with 100 interfaces. If a connection is attempted which exceeds the configured interface limit for a particular realm, the connection is refused.

The PIPE supports a few methods of administratively assigning network addresses and, where required, netmasks and forwarding prefixes (static routes), to the various FR, PPP and ATM link interfaces. In addition to the various link

interfaces the PIPE provides an abstracted "null" interface which can be used in conjunction with the forwarding function to provide for discard (or black-holing) of various categories of traffic. The appropriate methods are determined when a new interface is configured on the PIPE depending on the specific type of Access Interface/Service Interface/Core Interface required. Once an interface is defined, but before the configuration applied and it is activated, the interface is linked to the appropriate realm, ensuring that the traffic associated with that interface will only be forwarded within the correct network address spaces.

Typically, PPP links will either be configured using the "numbered-numbered" model, where the PPP peers are the only two nodes in a distinct point-to-point subnet, or the "unnumbered-unnumbered" model, where the peers have no IP addresses for the PPP interfaces on the PPP link. The link simply provides a bi-directional path between two distinct subnets. The PPP links may also be configured using the "numbered-unnumbered" model which means that only the interface address of the remote peer from PIPE is set for the link. For the "unnumbered-unnumbered" and the "numbered-unnumbered" models the PIPE supports the use of the "local route server" address to help manage control of these types of connections.

The local address assignments for ATM and Frame Relay service interfaces are provided from the Configuration Server/Route Server based on the PIPE providing the information required to determine which Service Interface/Access Interface is currently serviced by the PIPE.

Inverse ARP (InARP) is the standard method, in older, non-MPOA environments, for network devices to discover the IP

address of a peer device associated with a particular virtual circuit (e.g. ATM or Frame Relay). This allows for verification and dynamic configuration of address mappings rather than relying on static configuration of the ARP table. The PIPE can be configured to use InARP to discover the IP addresses of the network neighbors connected to the aggregate interfaces.. Some existing implementations of IP over NBMA media have no support for Inverse ARP. To allow interoperation, controls for disabling/enabling InARP and for static ARP table administration are provided via the PIPE management entities. Service interfaces established and configured using MPOA do not support InARP.

Address assignments for the "MPOA" ATM VC core interfaces are provided from the Configuration Server and Route Server The common controls for all Service Interfaces are Enabled/Disabled and Reset. In addition being able to disable, enable or reset the interface the operator can examine the state of the interface and view various interface statistics. There are many statistics and configuration details which are common to all interfaces. The PIPE provides all of the relevant values defined in the current IF MIB and also provides a number of useful summary statistics through various management interfaces. In addition diagnostics and controls have specific behaviors related to the various types of interfaces. Disable and Enable are used to temporarily block an interface from being used.

For PPP interfaces, Reset causes the PPP state machines to gracefully tear down the link and return to the initial state. This control is intended for forcing the controlled disconnection of specific PPP connections. For FR and ATM service interfaces, Reset causes the connect to redo any defined initial exchange. For both PPP and FR/ATM service

interfaces a reset causes all queues for the interface to be flushed.

Information relevant to tracing the PPP connection state is collected and made available through various management interfaces. Tracing of CHAP does not expose security specific details of the authentication protocol. The trace facility recognizes all assigned numbers for these PPP protocols listed in current IANA assigned numbers, including protocols and options not supported on the PIPE.

Information related to tracking the state of FR and ATM Service Interfaces and the ATM Core Interfaces is collected and made available through various management interfaces.

The PIPE provides a few control interfaces to aid in network and system diagnostics and maintenance:

1) Echo packet generation - provided to verify the IP protocol connectivity between PIPE and other network entities. The ICMP echo request is basis of the commonly used PING command. The PIPE can generate such requests and forward them to other network entities. The PIPE also replies to ICMP echo requests.

2) Network path tracing - provided for tracing the route IP traffic takes to reach a particular destination host interface. This function is equivalent with the "traceroute" command in UNIX. The mechanism involves launching a specially sequenced stream of UDP probe packets and then listening for ICMP time-exceeded (TTL-expired) responses from the forwarding devices along the path. The addresses of intermediate devices that responded as IP packets traversed the path are displayed along with an estimate of the delay based on the round trip for each transaction.

The PIPE supports Spanning Tree Protocol as defined in IEEE 802.1(d. The Spanning Tree implementation allows for loop-free topology such that a path exists between every pair of LANs in the network.

STP is negotiated on a per VPN basis, enabling each VPN to have a separate STP instance. STP does not apply to the Internet Access case.

Extensions to the standards are based on those defined below:

1) If the PIPE becomes unregistered all established SVCs are torn down, such that bridging traffic and STP BPDUs are not forwarded.

2) A configuration BPDU is recognized and ignored if it is received by its originator on the same port from which it was sent.

3) BPDU received over ATM from anything other than the through the "MPOA" client are ignored by the PIPE. (The '`MPOA" client will drop any BPDU that is not received from a registered device.)

4) If the Bridge Aggregate interface for the particular realm goes into a blocking state, the destination cache must be flushed to ensure that no entries point to the [now blocked] interface. In addition, when the Bridge Aggregate returns to the forwarding state, the source cache for the realm is flushed so that it can resynchronized with the MPOA client

5) Negotiation for the version of STP supported between registered devices is limited to protocol 1 (IEEE802.1 (d)) or NULL in the case where an external bridge does not support STP.

STP on the PIPE affects the state of one or more of its interfaces. Current STP states of the Service Interfaces are viewable via the NMTI management interface. The STP

standard, described in IEEE 802.1 (d), provides for the following configurable parameters:

Priority	used to determine the cost of using this bridge as root.
Max Age	amount of time before a configuration message should be deleted
Hello Time	time between configuration BPDU advertising root status
Forward Delay	length of time spend in intermediate state before changing from blocked to forward state
Aging Time	length of time since a root sent a configuration message

These parameters are configurable through a management interface and accessible via SNMP. Default STP parameters are used in the absence of user configured values.

The PIPE communicates with the Configuration Server to resolve which Route Server is controlling each of the Realms supported by the PIPE. The PIPE communicates with each Route Server to register and verify new service interfaces, to declare new locally attached hosts and subnets, and to resolve remote bridged or network-layer addresses to ATM addresses.

After being initialized from the Control Card, the PIPE first connects to the Configuration Server. It uses the address configured for the Configuration Server which defaults to a well-known AESA anycast address. The traffic parameters are configurable.

The PIPE will be downloaded with information about each VPN/IA/Realm within the system. This includes the ATM addresses of the primary and backup route servers.

As the information changes, the Configuration Server keeps each of the PIPEs updated.

The connection to the Configuration Server is maintained continuously using a persistent SVC. If the connection fails

or is released the persistent SVC mechanism will attempt a reconnect (with an initial period of 1 second) to the same anycast address and will continue to attempt the call indefinitely. Because of the nature of the anycast address mechanism when the new connection is eventually established it may even be to a different Configuration Server. The exact same procedures as explained for Initialization above apply to the new connection.

The Configuration Servers, in an N+1 redundant system of databases, distribute to each of the PIPEs the information necessary for establishing the LAN data and LAN control connections required for all the realms each of the PIPEs are serving

After receiving the ATM Addresses of all of the Route Servers, the PIPE establishes a LAN Data connection to each of the Route Servers for each of the VPN/IA/Realms that it has Service Interfaces for. The traffic parameters are configurable on a per-VPN/LA/Realm basis. The connection does not use any assured delivery capabilities.

When a Route Server detects a LAN Data connection having been established, the Route Server starts the registration mechanism by sending the Register Server message (i.e. supplies the features it supports) to the PIPE. The PIPE responds with a Register Client message (supplies the features the PIPE supports) back to the Route Server. The Route Server then sends a Register Response message which indicates a successful registration.

Following successful registration, the PIPE establishes a LAN Control to the Route Server. This connection uses different traffic parameters that are again configurable on a per-Realm basis, and using the Q.SAAL assured delivery bearer mechanism. This connection is used provide various elements of configuration information.

Also following successful registration, the Route Server will add the newly registered PIPE to a LAN Broadcast (point-to-multipoint) connection. The Route Server uses this connection for broadcast packets, multicast packets and for table downloads.

The LAN Data, LAN Control and LAN Broadcast connections are maintained continuously as long as Service Interfaces exist for the VPN. If a LAN Data or LAN Control connection is released the persistent SVC mechanism (with an initial period of 1 second) will attempt a reconnect using the current Route Server (e.g. primary) address. If the persistent SVC mechanism fails on the final exponential backoff to the current address, the PIPE clears any LAN Data, LAN Control and LAN Broadcast connections to the failed Route Server. An attempt is then made to set up the LAN Data connection to the other Route Server (e.g. backup) address, thereby restarting the registration process.

Since the PIPE cannot control it's addition to the LAN Broadcast connection, it cannot engage in the persistent SVC mechanism for this connection. Instead, the PIPE relies on the current (e.g. primary) Route Server to perform the persistent SVC mechanism. On detection of the loss of the LAN Broadcast connection the PIPE will however begin a timer of duration equivalent to, but slightly longer than the total duration of the persistent SVC mechanism's retry period. This timer is canceled should the errant LAN Broadcast connection be re-established. On expiry of this timer, the PIPE will clear any LAN Data or LAN Control connections to the failed Route Server. The PIPE will then attempt to set up the LAN Data connection to the other Route Server (e.g. backup) address, thereby restarting the registration process.

If the persistent SVC mechanism fails on the final exponential backoff to both Route Servers for a VPN/IA/Realm, then the PIPE informs the Configuration Server that that particular set of Route Servers is unreachable and a major alarm is raised on the 36170.

After ~1.3 times the Route Server cold-start time and including a random factor of +0.15 RS cold start time of outage of the LAN Data connection, the operation of this Realm ceases. All cache entries are removed. This limits the potential of creating forwarding loops and unintended black-holes within the network.

The PIPE supports bridged VLANs for any protocol family. Bridged VLANs separate traffic of different protocols and limit the protocols that can be used to communicate from specific hosts. They can carry all network-layer protocol families or any of the following:

1) IP
2) IPX (Internet Packet eXchange)
3) XNS (Xerox Network System)
4) SNA (Systems Network Architecture)
5) NetBIOS (Network Basic Input/Output System)
6) CLNP
7) Banyan VINES (Virtual Network System)
8) AppleTalk
9) DECnet
10) LAT (Local Area Transport)

VLAN membership is configured from the route server. There is no local support for configuring bridged VLANs..

The PIPE supports routed virtual subnets for the IP protocol only. Membership in a virtual subnet determines PPP IP address assignment, broadcast groups, etc.

Membership in virtual subnets is configured from the route server. There is no local support.

Service Interfaces can belong to multiple VLANs and Virtual Subnets. A Service Interface can belong to no more than one VLAN which supports the same protocol. A Service Interface can belong to many virtual subnets provided there is no overlap in assigned subnet IP addresses.

Except in the case of the Internet Access service, all other Realms (the VPNs) use the VIVID cache management protocols with the route server to learn and provide information about MAC and Network-layer addresses.

The Internet Access service uses Table Download (TD) in addition to the Cache Management protocols described above. The Table Download process begins with the Route Server providing the minimal set of cached Network-layer (IP) addresses required to allow the PIPE to begin processing. Following the initial table phase, the Table Download process continues with the final table phase. During this phase, the Route Server provides all remaining applicable Network-layer (IP) addresses.

At any time following the initial table download, table maintenance (adds & deletes) is performed using the VIVID cache management protocols described above.

Table Download may occur under any of three conditions:
1) Network cold start.
2) Partial network restart / cold start (multiple PIPEs).
3) Single PIPE restart / reconfig.
In fact, Table Download may begin under a single PIPE restart condition (3) which may later turn out to be a partial network restart condition (2). Table Download will utilize the unicast LAN Control SVC during the initial table

phase of Table Download. In order to provide good system start up performance without impacting the system when only a single PIPE is restarting, Table Download will utilize unicast (LAN Control) or multicast (LAN Broadcast) facilities depending on the number of PIPEs in the final table phase of Table Download. Table Download will also be capable of switching from using unicast (LAN Control) to multicast (LAN Broadcast) facilities as PIPEs enter the final table phase of Table Download.

Paths are constructed between Forwarders using SVCs set up using the ATM Address in the path table, the configured traffic descriptor for paths in the particular Realm, and B-HLI parameters indicating the type of device (the PIPE) that is establishing the connection. Parallel paths between Forwarders are disallowed except where difference levels of CoS are required. Two types of paths may be created between Forwarders (PIPEs)

1) aged; and
2) permanent.

The determination that a path is aged or permanent is made based on aging information provided by the Route Server when a path table entry (egress IP to ATM address mapping) is downloaded to the PIPE. The Route Server provides path table entries either as part of initial table download or on an exception basis.

Aged paths are set up on demand, whenever a datagram is received whose Network-layer (IP) address is mapped to an ATM Address where no SVC currently exists. These paths are aged out when there has been no data flowing over the connection for at configurable period of time. Age time is configurable on a per path basis. The default age time is 30 seconds. Aging out causes the SVC for the path to be released. When new data arrives for the path, the SVC is re-

established. While the path is being established or re-established, data is forwarded to the Default Forwarder

Permanent paths are set up as soon as a path table entry is provided to the PIPE by the Route Server and are maintained using the persistent SVC mechanisms. Should the persistent SVC for a path fail on its final exponential backoff, the Route Server will be informed so that routing information can be re-calculated. The PIPE will continue periodic attempts to re-establish the persistent SVC for the path. When the persistent SVC for the path is re-established, the Route Server is again notified so that that routing information can again be re-calculated

Paths may be viewed from a management interface. The paths the connections take through the network can only be derived manually. There is no call trace support for these connections.

N+M PIPE Redundancy is a form of warm redundancy that can optionally be enabled for the PIPE. The redundancy applies only within an individual 36170 and applies to the whole 36170. Separate independent N+M partitions are not available.

N PIPE cards are providing service to the N PIPE instances that have Service Interfaces programmed. M PIPE cards, referred to as the spare cards, are sitting around idle waiting for one of the N PIPE cards to fail.

A PIPE Instance is a floating set of functionality which can be placed on any PIPE Card within the 36170. It is identified by an 8-bit number. Service Interfaces are assigned to a PIPE Instance through management interaction. All CSI configuration, application maintenance, and statistics are performed by identifying the PIPE Instance,

not the PIPE slotId. The slotId is only used for card-specific maintenance, such as resetting, software downloading, etc. Everywhere else the PIPE instance is referred to as a PIPE.

The operation and the alarms that result from the operation of this redundancy scheme will be similar. The FS describes the dynamic nature of the assignment of PIPE Instances for service on PIPE Cards. It is to be noted that lower PIPE Instance numbers receive higher priority for assignment to a PIPE Card although the priority is non-preemptive.

When a non-spare (active) PIPE running applications becomes unavailable, all applications on the card are moved to a spare PIPE if it is available. Since PIPE N+M redundancy is not hot redundancy, the service interfaces and other applications are reset to the initial state. All current short-cuts and connections to the RS/CS are released. One of the formerly spare PIPE becomes active. This PIPE card starts setting up connections to the Configuration Server and the appropriate Route Servers and creates the necessary short-cuts.

Functions that are provided by the PIPE can be configured and managed through an external network management entity such as NMTI.
The major new managed entities provided on the PIPE are PIPE Instances, Realms, and Service Interfaces. Most of the configuration is downloaded to the PIPE by either the Configuration Server (CS) or the Route Server (RS), but can be displayed using NMTI. The configuration elements coming from the CS/RS arrive via a few different paths and methods dependent on the specific nature of the element. To distinguish these differences the following tags are provided:

- G - provided from the initial global configuration server

- X - generated as the result of an exception exchange with the RS

- W - direct write via basic configuration path

- D - derived from other configured element and the active state of the PIPE

The 36170 control card has no actual knowledge of Realms, and does not have any NVM storage allocated for the configuration of those entities.

The following table summarizes parameters that are configured and/or displayed for an entire 36170.

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
Switch Prefix	The 13-byte prefix MSN used for all PIPE Cards in the system. This is the Prefix assigned under CONFIG SVC SYSTEM ADDR_PREFIX ATM_END_ADDR	Null - This is the only required parameter for CSI besides the Service Interfaces.	R/W	-	R/W	-
Configuration Server ATM Address	The destination address of the Configuration Server. This is the Configuration Server that each PIPE will connect to when the PIPE is initializing.	The well-known AESA anycast address	R/W	-	R/W	-
Configuration Server Connection Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060 for details of stuff described under TRAFFIC menu (ABR not allowed)	UBR, PIR=155M, MIR=0.	R/W	-	R/W	-
PIPE N+1 Redundancy	On or Off	Off	R/W	-	R/W	-

Table 5-2.1 Internetworking 36170 System-Wide Configuration Table

The following table summarizes parameters that can be configured and/or displayed for a Realm.

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
Realm Name	16 character text string identifying the Realm.	None	R	-	-	W
Realm ID	5 digit number that globally identifies the Realm	None	R	-	-	W
Type	VPN, or Internet Access	VPN	R	-	-	W

Table 5-2.1 Internetworking Realm Configuration Table

The following table summarizes parameters that can be configured and/or displayed for a Realm configured on a specific PIPE instance (some of these parameters may be configured system-wide from the 46020, but they are downloaded to each PIPE card individually).

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
Local ATM Address	The address used to originate and terminate SVC call requests. This is a derived value, it is not configurable.	Switch CSI Prefix + PIPE MAC address	R	-	-	-
Local IP Address for the Realm	An IP address that can be used to "ping" the PIPE, or to originate ICMP echo, and traceroute requests from the PIPE.	Null	R	-	-	W
Number of Service Interfaces Configured for the Realm	16 Bit number - The total number of service interfaces that have been config-connected on the PIPE card for a particular realm.	0	R	-	-	D
Maximum SVCs allowed (short-cut + connections to RS)	To avoid having a particular realm steal all of the SVCs allowed per-PIPE, this provides a limit. The sum of the limits for each of the realms does not need to sum to the maximum per-PIPE.	Maximum number of SVCs per PIPE	R	-	-	W
Number of P2P Short-cut Paths	16 bit number - The number of Point to point short-cut paths for this realm from or to this PIPE Card. NOTE: This number is very dynamic in nature.	n/a	R	-	-	X
Number of P2MP Connection Leaves	16 bit number - The number of Point to multipoint paths from the MCS (or in the future, Edge Forwarders) for this realm to this PIPE Card. NOTE: This number is dynamic in nature.	n/a	R	-	-	X
Maximum ICMP Response Rate	8 bit number - The maximum number of ICMP messages that this PIPE card will generate for this realm over a 1 second period.	10	R	-	-	W
Flow Detection Parameters M and N	2 16 bit numbers - The number of packets (M) that must be sent within N seconds to a given IP address in order for a VPN to set up a short-cut path to that IP address.	4 pkts in 10 seconds	R	-	-	W
Shortcut SVC aging time	16 bit number - how long idle shortcut paths remain established	90 sec	R	-	-	W

Service I/F Connections to Access I/Fs	PVC or SPVC connections established by the 46020, and then bound to this Realm via the CS/RS (see § 5.2.6).	Null Realm (all traffic is dropped)	R	-	-	W

Table 5-2.2 Internetworking Realm PIPE Configuration Table

The following tables summarize parameters that are configured and/or displayed for Bridged VLANs on a PIPE card.

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
PIPE Adapter ID	32 bit number that the RS uses to uniquely identify the PIPE within the Realm.	n/a	R	-	-	W
Spanning Tree Administrative Status	Enabled or Disabled	Enabled	R	R	-	W
Spanning Tree Bridge Priority	1..TBD	1	R	R	-	W
Spanning Tree Maximum Age	0.1..TBD.0 seconds	20.0	R	R	-	W
Spanning Tree Hello Time	0.1..TBD.0 seconds	2.0	R	R	-	W
Spanning Tree Forwarding Delay	0.1..TBD.0 seconds	15.0	R	R	-	W
Spanning Tree Root Path Cost	1...TBD	TBD	R	R	-	W
Spanning Tree Root Port	The service interface (or the ATM aggregate) leading to the root of the Spanning Tree.	ATM	R	R	-	W
Spanning Tree Port Status	Disabled, Blocking, Learning, Forwarding	n/a	R	R	-	D
Local Address Table	List of MAC addresses reachable from each Service Interface in the Realm (from this PIPE card).	n/a	R	R	-	X
Remote Address Table	List of MAC addresses reachable in the Realm over the ATM aggregate, and the ATM addresses they can be reached at (from this PIPE card).	n/a	R	R	-	X

Table 5-2.1 Realm Bridging Parameters

Parameter	Description	Default	MIT	SNMP	NGT	RS/ CS
VLAN Id	12 bit number - Unique number per-Realm (this is used as the key for management operations)		R	-	-	W
Route Server Address	Any ATM Address Format	This address is given to the PIPE Card by the Config Server for establishing SVCs to that Route Server	R	-	-	G
Broadcast Server Address	Any ATM Address Format	This address is given to the PIPE Card by the RS for establishing SVCs to the Broadcast Server	R	-	-	W
Multicast Server Address	Any ATM Address Format	This address is given to the PIPE Card by the RS for establishing SVCs to the MCS	R	-	-	W
Default Forwarder Address	Any ATM Address Format	This address is given to the PIPE Card by the Route Server for establishing SVCs to the Default Forwarder	R	-	-	W
Route Server LAN Data Connection Status	Up / Down	The status of the SVC for LAN Data	R	-	-	-
MCS LAN Data Connection Status	Up / Down	The status of the SVC for LAN Data	R	-	-	-

Route Server Broadcast Connection Status	Up / Down	The status of the SVC for Broadcasts from the Broadcast Server	R	-	-	-
Route Server LAN Data Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
Route Server LAN Control Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
Broadcast Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
Routed Protocols	IP	Disabled	R	-	-	W
Bridged Protocols	A bitmask indicating state for: IP, IPX, XNS, SNA, NetBIOS, CLNP, VINES, AppleTalk, DECnet, LAT, other	Disabled	R	-	-	W
Service Interfaces	A Bitmask of 700 service interfaces, that get mapped to VPI/VCI 0/101 through 0/800 - Displayed in NMTI as VPI/VCI. A "1" in the mask indicates that the Service Interface belongs to the VLAN.	0	R	-	-	W

Table 5-2.1 VLAN Bridging Parameters

The following tables summarize parameters that are configured and/or displayed for IP Routed Virtual Subnets on a PIPE card.

Parameter	Description	Default	[illegible]	SNMP	NCT	RS/CS
PIPE ID	32 bit number that the RS uses to uniquely identify the PIPE within the Realm.	n/a	R	-	-	W
IP Forwarding Table	List of IP addresses reachable within the Realm (from this PIPE card), and the next-hop IP address and service interface (or ATM aggregate) over which they can be reached	n/a	R	R	-	X
Local ARP Table	List of IP addresses reachable from each Service Interface in the Realm (on this PIPE card), and their associated MAC address, if one exists	n/a	R	R	-	X
Remote ARP Table	List of IP addresses reachable in the Realm over the ATM aggregate (for this PIPE card), and the ATM addresses they can be reached at	n/a	R	R	-	X

Table 5-2.1 Realm Routing Parameters

Parameter	Description	Default	NMU	SNMP	NCI	RS/CS
Subnet Id	12 bit number - Unique number per-Realm (this is used as the key for management operations)		R	-	-	W
Route Server Address	Any ATM Address Format	This address is given to the PIPE Card by the Config Server for establishing SVCs to that Route Server	R	-	-	G
Multicast Server Addresses	Any ATM Address Format	The addresses given to the PIPE Card by the RS for establishing SVCs to Multicast Servers	R	-	-	W
Broadcast Server Address	Any ATM Address Format	This address is given to the PIPE Card by the Route Server for establishing SVCs to the Broadcast Server	R	-	-	W
Default Forwarder Address	Any ATM Address Format	This address is given to the PIPE Card by the Config Server for establishing SVCs to the Default Forwarder	R	-	-	W
Route Server LAN Data Connection Status	Up / Down	The status of the SVC for LAN Data	R	-	-	-
Broadcast Server LAN Data Connection Status	Up / Down	The status of the SVC for data to the Broadcast Server	R	-	-	-

MCS LAN Data Connections Status	Up / Down	The status of the SVCs for data to the MCS	R	-	-	-
Route Server LAN Control Connection Status	Up / Down	The status of the SVC for LAN Control	R	-	-	-
Broadcast Server Connection Status	Up / Down	The status of the SVC for Broadcasts from the Broadcast Server	R	-	-	-
Multicast Server Connection Status	Up / Down	The status of the SVC for Multicasts from the MCS	R	-	-	-
Route Server LAN Data Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
Route Server LAN Control Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
Multicast/Broadcast Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060.	The parameters that are in effect at this time	R	-	-	W
IP Address (Route Server Interface)	IP Address - The address of the route server interface on the virtual subnet.		R	-	-	W
IP Address Prefix Length	8 bit number ranging from 1...32. The significant portion of the IP Address referenced above for IP addresses within the subnet		R	-	-	W
IP Broadcast Address	IP Address - The IP Address for Subnet Broadcasts		-	-	-	W
MTU	number between 100 and 9182	1500	R	-	-	W

Service Interfaces	A Bitmask of 700 service interfaces, that get mapped to VPI/VCI 0/101 through 0/800 - Displayed in NMTI as VPI/VCI. A "1" in the mask indicates the service interface is part of the subnet.	0	R	-	-	W
Admin Status	UP/DOWN	Current administrative state of the VLAN	-	R	-	W

Table 5-2.1 IP Routed Virtual Subnet Parameters

The following tables summarize parameters that are configured and/or displayed for Service Interfaces on a PIPE card.

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
Realm Name	16 character text string identifying the Realm that the Service Interface has been associated with. A Service Interface can only belong to 1 realm.	None	R	-	-	W
Realm ID	5 digit number that globally identifies the Realm that the Service Interface has been associated with. A Service Interface can only belong to 1 realm.	None	R	-	-	W
Type	The type of Realm the Service interface is associated with - either VPN, or Internet Access	VPN	R	-	-	W
Access Interface	The Shelf-Slot-Port-<logical channel> that the Service interface is connected to within this switch.	None	R/W	-	R/W	-
Traffic Parameters	See 31FS0058, 31FS0059, and 31FS0060. The traffic parameters for the connection between the Service Interface and the Access Interface.	None	R/W	-	R/W	-
Protocol	Protocol that the Access Interface or Service Interface is terminating. PPP, RFC1483, or RFC1490.	None	R	-	-	W
Status	Up, or Down. A status of Up indicates that the physical layer, and all control protocols are up on the service interface.	Down	R	-	-	-

Table 5-2.1 Service Interface Parameters

In addition to the general Service Interface Configuration shown above, the following tables summarize additional parameters that are configured and/or displayed for PPP Service Interfaces on a PIPE card.

Parameter	Description	Default	NMTI	SNMP	NGT	RS/CS
LCP Link Status	Disabled/Down/Negotiating/Up	Down	R	-	-	W/D
MRU	The largest PPP frame the PIPE is willing to receive on this service interface	1500	R	-	-	W
MTU	The largest PPP frame the PIPE will transmit on this service interface	1500	R	-	-	W
Magic Number	A magic number	0	R	-	-	W
Protocol Field Compression	Enabled/Disabled	Enabled	R	-	-	W
Address & Control Field Compression	Enabled/Disabled	Enabled	R	-	-	W
Identification	PIPE 90 # plus SW release		R	-	-	-
Link Mode	Active/Passive/Silent - Active keeps trying to establish up to the configure retry limits. In passive mode, one config-request is sent - if no reply is received, then wait for a config-request from the peer. In silent mode, the service interface just waits for a config-request from the peer.	Passive	R	-	-	W
Echo Interval	16 Bit number (max 3600). Sends an LCP echo-request frame to the peer every n seconds, as specified by this parameter. 0 disables.	0	R	-	-	W
Echo Failure	8 bit number, number of consecutive LCP echos sent with no reply before the PPP connection is terminated.	3	R	-	-	W
Max Configure	8 bit number, maximum number of LCP config-request transmissions.	10	R	-	-	W
Max Failure	8 bit number - maximum number of LCP config-NAKs returned before starting to send config-rejects instead	10	R	-	-	W

Max Terminate	8 bit number - Maximum number of LCP terminate-request transmissions.	3	R	-	-	W
Restart Interval	8 bit number - LCP retransmission timeout	3	R	-	-	W

Table 5-2.1 PPP Service Interface LCP Parameters

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
CHAP Link Status	Disabled/Down/Negotiating/Up	Disabled	R	-	-	W/D
CHAP Interval	16 bit number, rechallenge interval in seconds (0 to never rechallenge)	0	R	-	-	W
CHAP Max Challenge Count	8 bit number - maximum number of CHAP challenge transmissions	10	R	-	-	W
CHAP restart interval	8 bit number - retransmission timeout (in seconds) for challenges	3	R	-	-	W
CHAP Local Secret	Something secret	None	-	-	-	W
CHAP Peer Secret	Something secret	None	-	-	-	W
CHAP Local Name		None	R	-	-	W
CHAP Remote Name		None	R	-	-	W

Table 5-2.1 PPP Service Interface CHAP Parameters

Parameter	Description	Default	NMTI	SNMP	NCI	RS/CS
BCP Link Status	Disabled/Down/Negotiating/Up	Disabled	R	-	-	W/D
Tiny-gram compression	Enabled/Disabled	Disabled	R	-	-	W
Mac-Address		Null	R	-	-	W
STP Negotiation	802.3 only, Enabled/Disabled	Enabled	R	-	-	W

Table 5-2.1 PPP Service Interface BCP Parameters

Figure 5.a shows the top-level NMTI Softkey Legend for "CONFIG". The corresponding sub-menus are shown in the sections that follow.

Most of the information shown on these menus is downloaded from either the Config Server or the Route Server. In fact, the only items that can be modified via NMTI are the items that appear under the CONFIG_SERVER softkey.

The NMTI menus go several levels deep on many screens. In many cases, the complete chain of commands will not fit on a single command line. In those circumstances, the word "CONFIG/MAINT/STATS" remains on the command line, but words immediately to the right of the word "CONFIG/MAINT/STATS" may be deleted to make room for the most recent softkey displays. In those instances, the deleted text is restored as the "CANCEL" softkey is used to back out of the lower menus.

<u>NMTI Softkey Legend</u>
Existing Functionality
New Functionality
«keystroke input»
Toggle keys are defined twice at same level (default value is on top)
F1: CONFIG
    F7: MORE
        F1: INTERNETWORK
            F1: PIPE_INSTANCE
            F2: CONFIG_SERVER
            F4: PIPE_REDUND

Figure 5.a General Softkey Infrastructure.

Virtually all the configuration information the PIPE card receives from the Configuration Server and the Route Server can be viewed from the PIPE_INSTANCE NMTI screens. The PIPE_INSTANCE menu tree is shown in Figure 12. The SERVICE_I/F, BRIDGING, and IP_ROUTING subtrees are expanded upon later in this document. Note that the "BRIDGING" softkey is only available if the Realm has been configured as a VPN. If the Realm has been configured to provide Internet Access, the error message "Bridging Only Supported For VPN's" will be displayed when the "BRIDGING" softkey is selected.

It should be noted that the "SERVICE_~/F ' softkey directly below <pipeInstance> is simply a shortcut to "PIPE_INSTANCE <pipeId> REALM <realmId> SERVICE_I/F'
While particular embodiments of the invention have been described and illustrated it will be apparent to one skilled in the art that numerous changes can be made to the basic concept. It is to be understood that such changes will fall within the full scope of the invention as defined by the appended claims.

CLAIMS

1. In a system for delivering internetworking service functions utilizing internetworking devices to provide said services to two or more specific network users, said method comprising:
logically partitioning said devices into sub-elements;
allocating said sub-elements to independent realms; and
assigning said independent realms to said specific network users.

2. A method as defined in claim 1 wherein each of said realms is a specific instance of an internetworking service function.

3. A method as defined in claim 2 wherein said specific instance is a public Internet access service.

4. A method as defined in claim 2 wherein said specific instance is a virtual private network (VPN) service.

5. A method as defined in claim 4 wherein said VON service is a bridged connectivity service.

6. A method as defined in claim 4 wherein said VPN service is a network layer connectivity.

7. A method as defined in claim 1 wherein said internetworking devices include an ATM backplane.

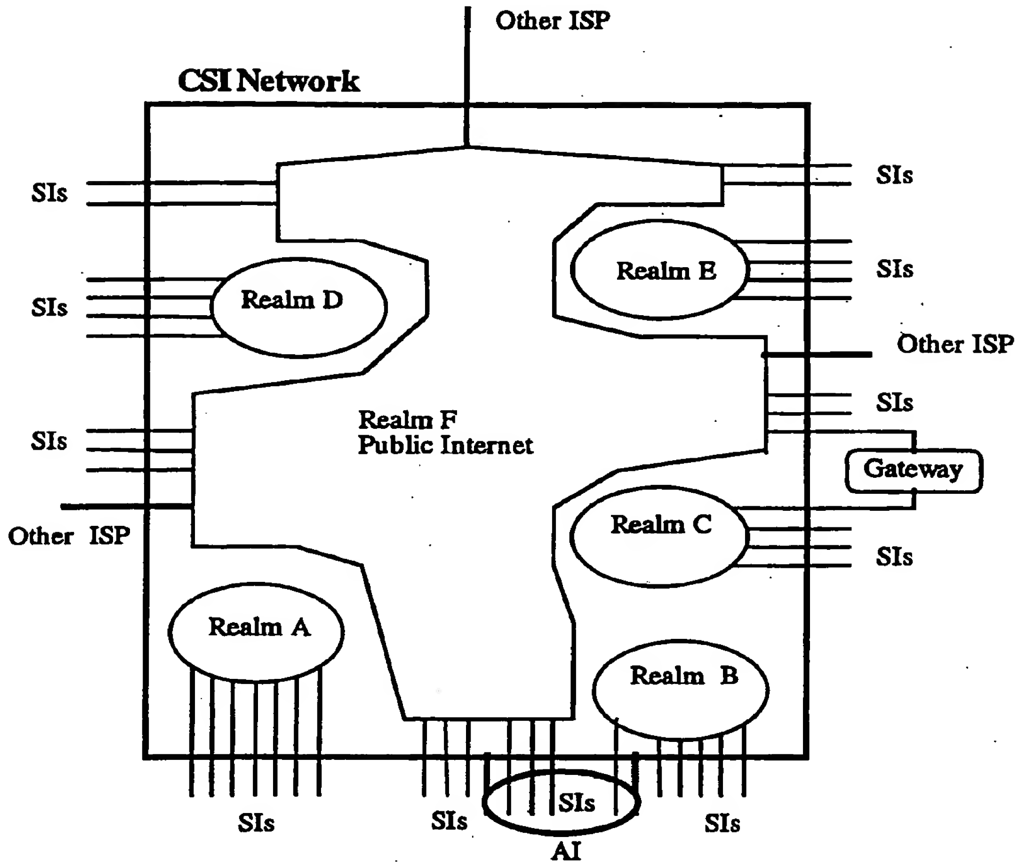

FIGURE 1

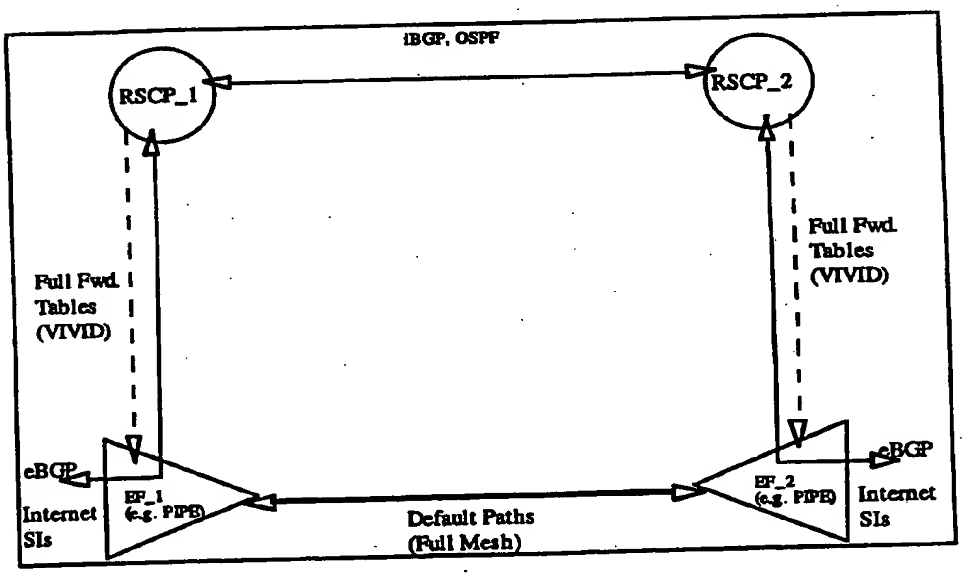

FIGURE 3

Marks & Clerk

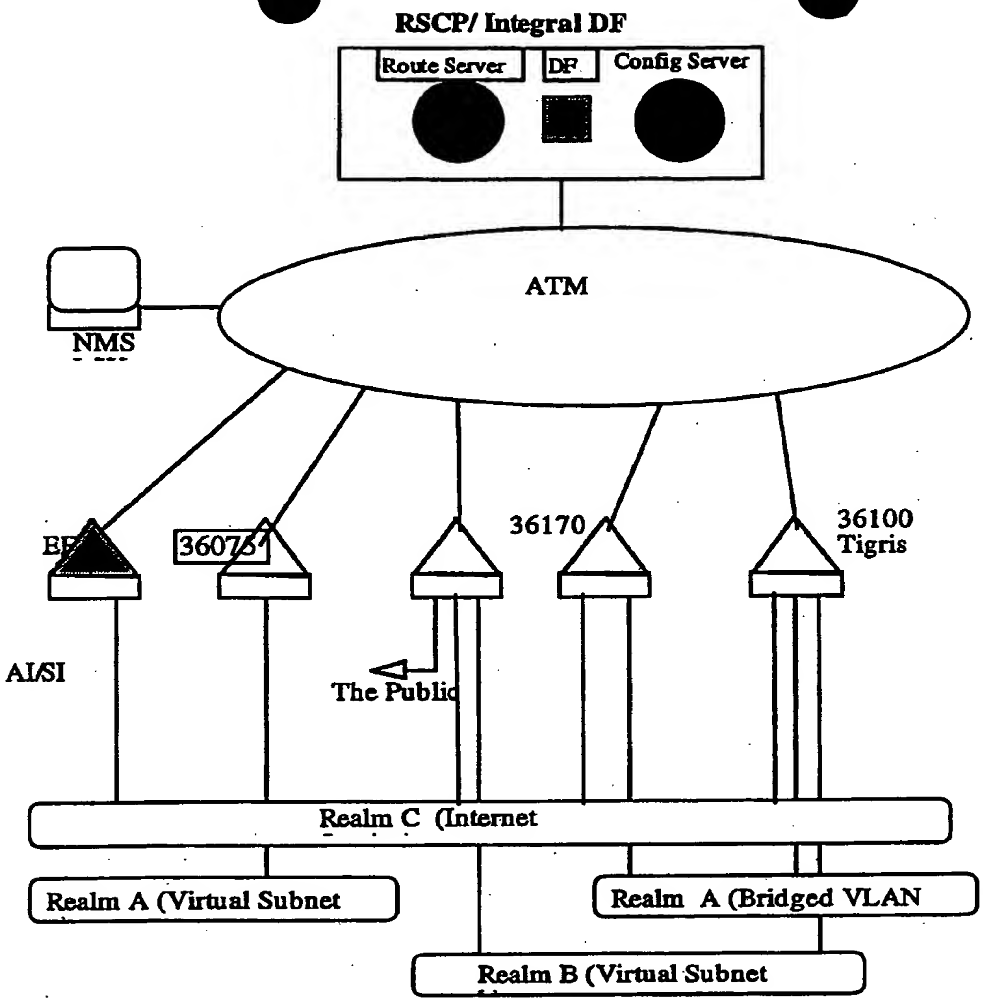

FIGURE 2

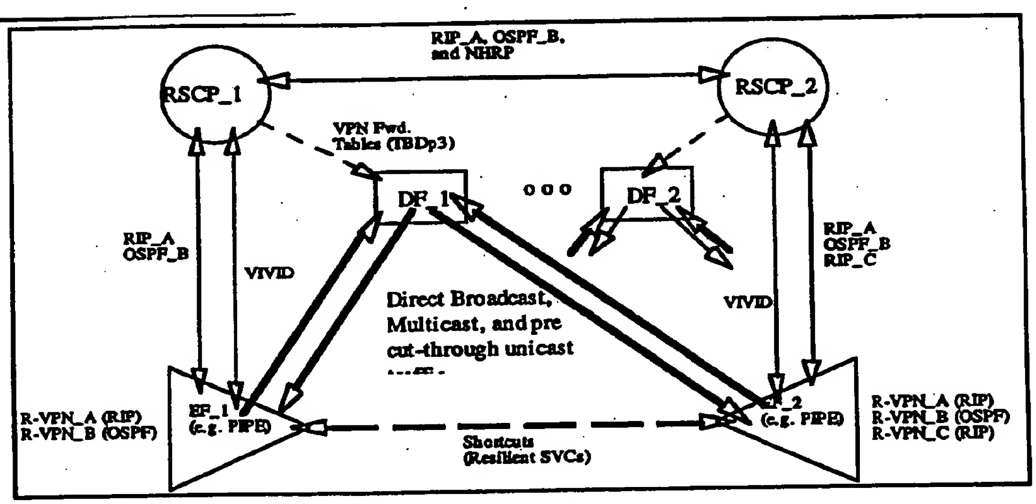

FIGURE 4

Marks & Clerk

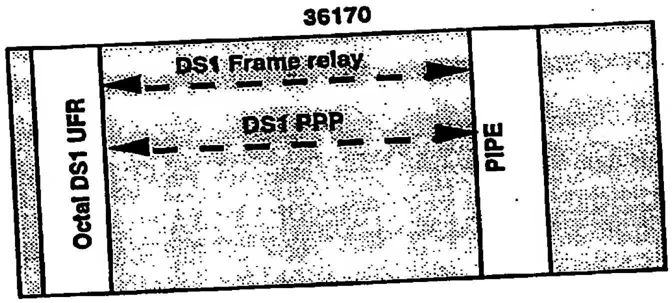

FIGURE 5

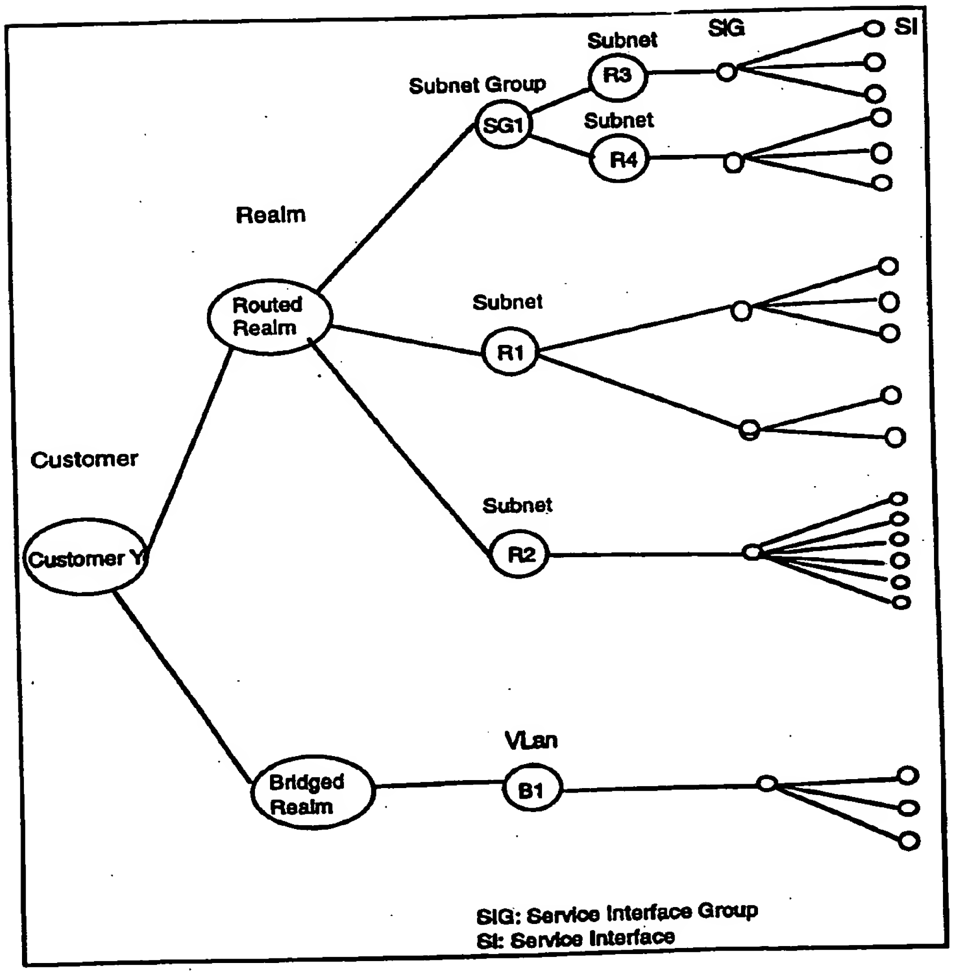

FIGURE 8

Marks & Clerk

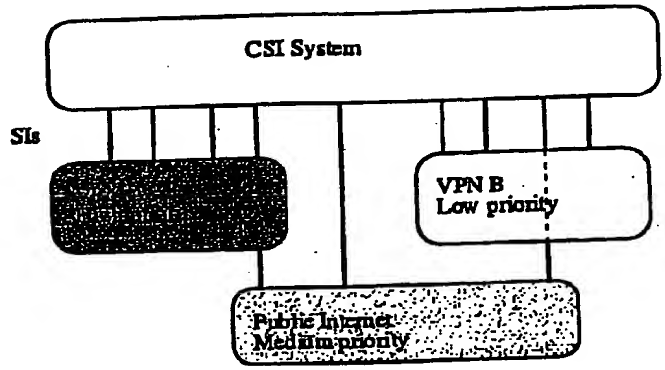

FIGURE 6

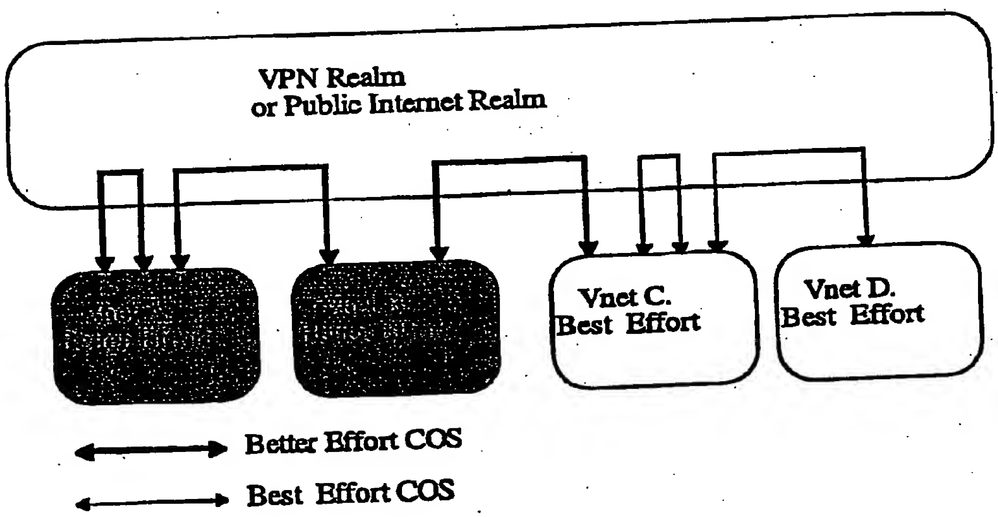

FIGURE 7

Marks & Clerk

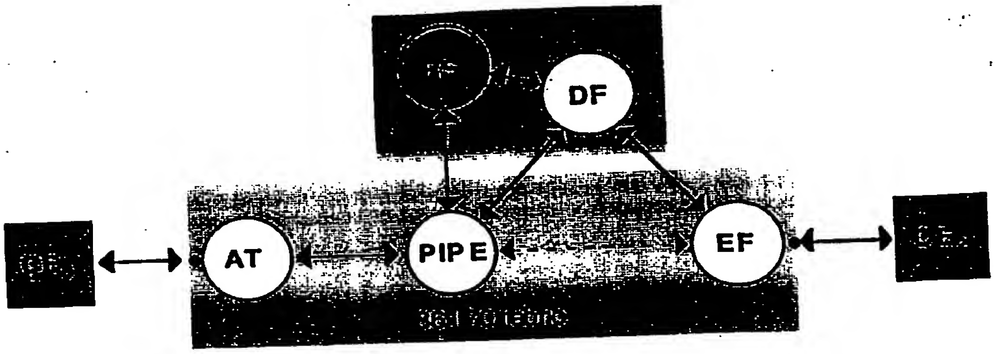

FIGURE 9

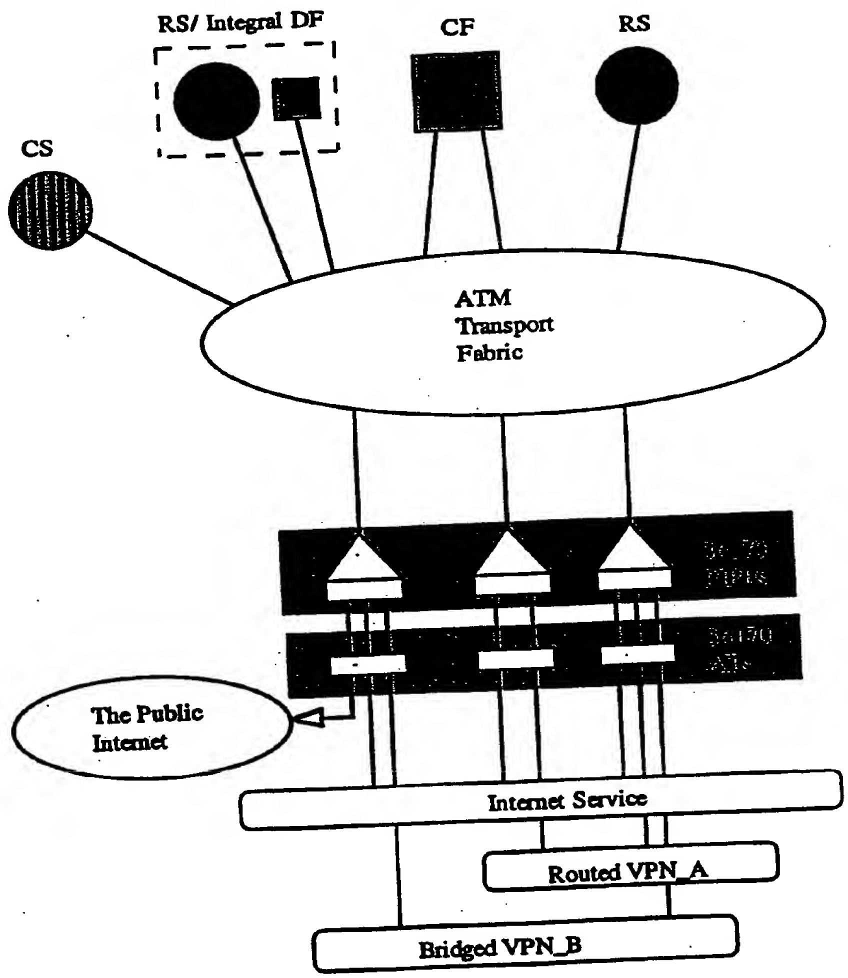

FIGURE 10

Marks & Clerk

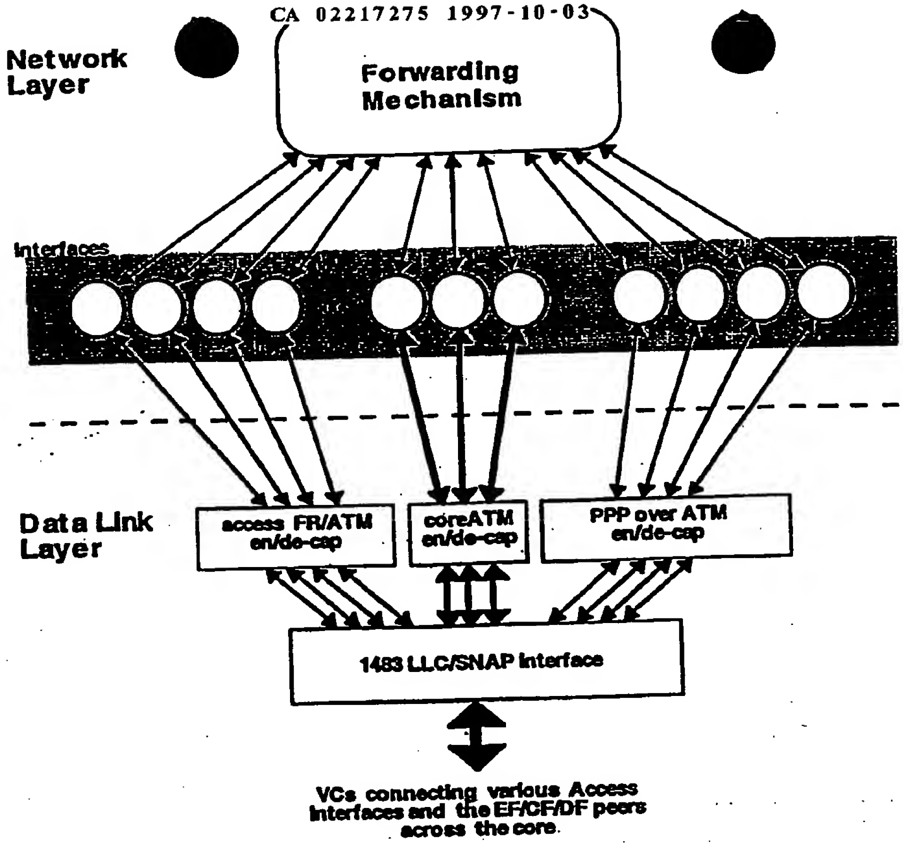

FIGURE 11

```
          1         2         3         4         5         6         7         8
 12345678901234567890123456789012345678901234567890123456789012345678901234567890
 ------------------------ product specific header line -------------------------

 PIPE Instance: PIPE-01     PIPE Name: MY_PIPE             Active PIPE Card: P2-8

 Number of Service Interfaces configured on this PIPE:  400
 Number of P2P short-cut paths active on this PIPE:     650
 Number of P2MP connection leaves active on this PIPE:    8
 Configuration Server Link Status:  Up

 Realm   Realm Name         Realm Type          Serv. I/Fs    P2P SVCs    P2MP SVCs
 -----   ----------------   ----------------    ----------    --------    ---------
   1     Newbridge Ntwks    VPN                      75          200          3
   2     Crosskeys          VPN                     120          150          4
  25     Joe's ISP          Internet Access         205          300          1





  CONFIG MORE INTERNETWORK PIPE_INSTANCE PIPE-01


  1-REALM          2-PAGE_DOWN      3-               4-               5-SERVICE_I/F
  6-PAGE_UP        7-               8-CANCEL         9-QUIT           0-
  0                1                2                3                4
  0123456789ABCDEF0123456789ABCDEF0123456789ABCDEF0123456789ABCDEF0123456789ABCDEF
```

FIGURE 12

Marks & Clerk